



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 180127

TO: Phuong Bui

Location: REM-2a!5&2C15

Art Unit: 1638

Tuesday, August 24, 2004

Case Serial Number: 09/900237

From: David Schreiber

Location: Biotech-Chem Library

Remsen E01A61

Phone: 272-2526

david.schreiber@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:11:20 ; Search time 13694 Seconds
(without alignments)
11476.694 Million cell updates/sec

Title: US-09-900-237a-29
Perfect score: 3626
Sequence: 1 gcaaggaagaccgcctcca.....tcacatttgaggaggtttt 3626

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3626	100.0	3626	8	BT009438	BT009438 Triticum
2	2694.4	74.3	3954	8	AK072356	AK072356 Oryza sat
3	2637.8	72.7	3264	6	AX653232	AX653232 Sequence
4	2595	71.6	3795	8	AF200533	AF200533 Zea mays
5	2595	71.6	3799	6	AX338680	AX338680 Sequence
6	2572.2	70.9	3745	8	AF200528	AF200528 Zea mays
7	2360.8	65.1	4282	8	AK069196	AK069196 Oryza sat
8	2352.4	64.9	3222	6	AX652952	AX652952 Sequence
9	2325.2	64.1	3676	8	AF200529	AF200529 Zea mays
10	1720	47.4	3532	8	AY055724	AY055724 Populus t
11	1699.8	46.9	3723	8	AF150630	AF150630 Gossypium
12	1679.4	46.3	3229	8	BT002335	BT002335 Arabidops
13	1675.2	46.2	3682	8	AF027174	AF027174 Arabidops
14	1673.6	46.2	3614	6	AR267559	AR267559 Sequence
15	1673.6	46.2	3614	6	AX030946	AX030946 Sequence
16	1673.6	46.2	3614	6	BD022678	BD022678 Manipulat
17	1412.4	39.0	3732	8	AK099228	AK099228 Oryza sat
18	1412.4	39.0	3768	8	AK102140	AK102140 Oryza sat
19	1412.4	39.0	3801	8	AK099281	AK099281 Oryza sat
20	1412.4	39.0	3802	8	AK067967	AK067967 Oryza sat
21	1410.8	38.9	3897	8	AK100188	AK100188 Oryza sat
22	1409.8	38.9	3764	8	AK098978	AK098978 Oryza sat
23	1408.4	38.8	3725	8	AF200526	AF200526 Zea mays
24	1383.6	38.2	3851	6	BD236020	BD236020 Materials
25	1373	37.9	3752	8	AF200525	AF200525 Zea mays
26	1369.4	37.8	3640	8	AY162181	AY162181 Populus t
27	1343.4	37.0	3603	6	AX030942	AX030942 Sequence
28	1343.4	37.0	3603	6	BD022676	BD022676 Manipulat
29	1343.4	37.0	3763	8	BT008654	BT008654 Arabidops
30	1343.4	37.0	3763	8	BT008654	BT008654 Arabidops
31	1341.8	37.0	3673	6	AR267560	AR267560 Sequence
32	1341.8	37.0	3673	6	AX030948	AX030948 Sequence
33	1341.8	37.0	3673	6	BD022679	BD022679 Manipulat
34	1340.2	37.0	3812	8	AF200532	AF200532 Zea mays
35	1331.4	36.7	2828	8	AF200527	AF200527 Zea mays
36	1331.4	36.7	2830	6	AX338676	AX338676 Sequence
37	1316.8	36.3	3538	8	AF200530	AF200530 Zea mays
38	1314.6	36.3	3746	8	AK100914	AK100914 Oryza sat
39	1294.6	35.7	4127	8	AK121193	AK121193 Oryza sat
40	1281.2	35.3	4029	8	AK100877	AK100877 Oryza sat
41	1280.4	35.3	4208	8	AK073561	AK073561 Oryza sat
42	1271	35.1	3968	8	AF200531	AF200531 Zea mays
43	1246.2	34.4	3631	8	AK121170	AK121170 Oryza sat
44	1244.8	34.3	3847	8	AF304374	AF304374 Nicotiana
45	1235.6	34.1	3221	8	AK067386	AK067386 Oryza sat

ALIGNMENTS

RESULT 1
BT009438
LOCUS
DEFINITION Triticum aestivum clone wlmk4.pk0015.all:fls, full insert mRNA
3626 bp mRNA linear PLN 20-JUN-2003

ACCESSION BT009438
VERSION BT009438.1
KEYWORDS GI:32128989
SOURCE Triticum aestivum (bread wheat)
FLI CDNA.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 3626)
Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,

Pred. No. is the number of results predicted by chance to have a

Caraher, N.R., Hanafey, M.K. and Hainey, C.F.

TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

FEATURES

source

Location/Qualifiers
1..3626
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wlmk4.pk0015.all.fis"

ORIGIN

Query Match 100.0%; Score 3626; DB 8; Length 3626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACAGGAAACCCGCTCCAGCTCTGTCGTCGGTGGATCGCTGCGCGCCA 60
Db 1 GCACAGGAAACCCGCTCCAGCTCTGTCGTCGGTGGATCGCTGCGCGCCA 60

QY 61 TGAACGGCAGCGGACCGCTGAAGTCCGGGAGGCA CGGGCCGGGAGCTGTGCAGA 120
Db 61 TGAACGGCAGCGGACCGCTGAAGTCCGGGAGGCA CGGGCCGGGAGCTGTGCAGA 120

QY 121 TCTGCGCCGACGCGCTGGGACCACTTGAACGGGACGCTTTCACGGCTCGAGCTCT 180
Db 121 TCTGCGCCGACGCGCTGGGACCACTTGAACGGGACGCTTTCACGGCTCGAGCTCT 180

QY 181 GCCGCTTCCCGCTGTCGCGCCCTGCTTACGAGCAGCGCAGAGGGGACCCAGGCTT 240
Db 181 GCCGCTTCCCGCTGTCGCGCCCTGCTTACGAGCAGCGCAGAGGGGACCCAGGCTT 240

QY 241 GCCTCCAGTCAAGCAAGTACAAGCCACAGAGGAGCCAGCGATCCGCGGGAGG 300
Db 241 GCCTCCAGTCAAGCAAGTACAAGCCACAGAGGAGCCAGCGATCCGCGGGAGG 300

QY 301 AAGCGCAGCACTGATGCGCGATGATGATGATGATGATGATGATGATGATGATG 360
Db 301 AAGCGCAGCACTGATGCGCGATGATGATGATGATGATGATGATGATGATGATG 360

QY 361 AGGACAGAGCAGAGATGCTGACAGGATGCGCAGCTGCGCGATGAACACCGGGGCA 420
Db 361 AGGACAGAGCAGAGATGCTGACAGGATGCGCAGCTGCGCGATGAACACCGGGGCA 420

QY 421 GTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCTCTCAAGTATGACA 480
Db 421 GTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCTCTCAAGTATGACA 480

QY 481 GTGGAGATGCTTAGGGGATAGTCCCTTCAGTCAACACAGCCAGATGTCAGGAGAAA 540
Db 481 GTGGAGATGCTTAGGGGATAGTCCCTTCAGTCAACACAGCCAGATGTCAGGAGAAA 540

QY 541 TCCTCGAGCTTCGCTGATCATCATGATGTCCTTACGGGAAACATCAGCAGCGTG 600
Db 541 TCCTCGAGCTTCGCTGATCATCATGATGTCCTTACGGGAAACATCAGCAGCGTG 600

QY 601 CTCGCTTCCCTATGTAATCATTACCAATTCGTCAGGGATTCCTCGGCAGTATTG 660
Db 601 CTCGCTTCCCTATGTAATCATTACCAATTCGTCAGGGATTCCTCGGCAGTATTG 660

QY 661 GGAATGTCCTGGAAGAGAGAGTTCATGCTGGAATAAGCAGGAGTTCAGGAGTTCGGA 720
Db 661 GGAATGTCCTGGAAGAGAGAGTTCATGCTGGAATAAGCAGGAGTTCAGGAGTTCGGA 720

QY 721 TTCCATGACTAATGGACAAAGCATTCCTTCTGAAGTTCGGGAGCTACTGACATCG 780
Db 721 TTCCATGACTAATGGACAAAGCATTCCTTCTGAAGTTCGGGAGCTACTGACATCG 780

QY 781 ATGCATCTACTGAATACAAATGGAAGCGCTTTACTGAATGATGAACCTCGCCAGCTC 840
Db 781 ATGCATCTACTGAATACAAATGGAAGCGCTTTACTGAATGATGAACCTCGCCAGCTC 840

QY 841 TATCTAGAAAAGTCCCAATGCTTCTCTCAAAATAAATCCCTACAGAATGGTCAATGTTTC 900
Db 841 TATCTAGAAAAGTCCCAATGCTTCTCTCAAAATAAATCCCTACAGAATGGTCAATGTTTC 900

QY 901 TCGCGTTGGTTGTTCTAAGCATCTTCTCGCATCTACCGTCTCAAAATCTCTGCGTAAATG 960
Db 901 TCGCGTTGGTTGTTCTAAGCATCTTCTCGCATCTACCGTCTCAAAATCTCTGCGTAAATG 960

QY 961 CATACCCACTGGCTTTTATCTGTATATGTAGATTTGGTTGCTTTATCTCTGATAC 1020
Db 961 CATACCCACTGGCTTTTATCTGTATATGTAGATTTGGTTGCTTTATCTCTGATAC 1020

QY 1021 TGGATCAGTTCCCGAAGTGGTTTCCAATCAACCGGAGACCTACCTTGTATAGACTGGCTT 1080
Db 1021 TGGATCAGTTCCCGAAGTGGTTTCCAATCAACCGGAGACCTACCTTGTATAGACTGGCTT 1080

QY 1081 TAAGGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTTGTGAGTA 1140
Db 1081 TAAGGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTTGTGAGTA 1140

QY 1141 CAGTCCAGCCCTTGAAGGAGCCACCTATCTGCTCACTGCCAACACTGTGCTATCCATTCTTG 1200
Db 1141 CAGTCCAGCCCTTGAAGGAGCCACCTATCTGCTCACTGCCAACACTGTGCTATCCATTCTTG 1200

QY 1201 CTGTTGATTTATCCCGTGGACAAAGTCTCTTGTCTATCTATCTGATGACGAGCTTCAATGC 1260
Db 1201 CTGTTGATTTATCCCGTGGACAAAGTCTCTTGTCTATCTATCTGATGACGAGCTTCAATGC 1260

QY 1261 TGAATTTTGAAGGATGAGCTTGAAGGAGCCACCTATCTGCTCACTGCCAACACTGTGCTATCCATTCTTG 1320
Db 1261 TGAATTTTGAAGGATGAGCTTGAAGGAGCCACCTATCTGCTCACTGCCAACACTGTGCTATCCATTCTTG 1320

QY 1321 AGAAGTATGACATTTGAACCCAGAGCTCCGAGTTTCTTCTTCCAGAAAATGATTTACC 1380
Db 1321 AGAAGTATGACATTTGAACCCAGAGCTCCGAGTTTCTTCTTCCAGAAAATGATTTACC 1380

QY 1381 TGAAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGCGCATGAAGAGAGATATG 1440
Db 1381 TGAAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGCGCATGAAGAGAGATATG 1440

QY 1441 AGAATTTAAATCAGATTAATGCGCTAGTTTCTAAGCATTTGAAGTCCCGAGGAG 1500
Db 1441 AGAATTTAAATCAGATTAATGCGCTAGTTTCTAAGCATTTGAAGTCCCGAGGAG 1500

QY 1501 GATGGATCATGCAAGATGGCACCATGGCCAGAAAACATACAGGGATCATCTCTGAA 1560
Db 1501 GATGGATCATGCAAGATGGCACCATGGCCAGAAAACATACAGGGATCATCTCTGAA 1560

QY 1561 TGATTCAGGTTTCTTGGTCAAGTGGTGGCTTGTATCTGAGGGTATGAGCTCCCTCC 1620
Db 1561 TGATTCAGGTTTCTTGGTCAAGTGGTGGCTTGTATCTGAGGGTATGAGCTCCCTCC 1620

QY 1621 GTTTAGTTATGCTCTCGTGAAGAGCTCTGGTTCCAGCACCACCAAGAGCTGGTG 1680
Db 1621 GTTTAGTTATGCTCTCGTGAAGAGCTCTGGTTCCAGCACCACCAAGAGCTGGTG 1680

QY 1681 CCATGAATGCCCTTGTTCGTGCTCTCAGCTGCTCTTACTAATGGACATATCATGTTGATC 1740
Db 1681 CCATGAATGCCCTTGTTCGTGCTCTCAGCTGCTCTTACTAATGGACATATCATGTTGATC 1740

QY 1741 TTGATTTGATCATCATCAACAGCAGGCTGTCCGAGAGCTATGTGCTTCTCTAA 1800
Db 1741 TTGATTTGATCATCATCAACAGCAGGCTGTCCGAGAGCTATGTGCTTCTCTAA 1800

QY 1801 TGGATCAAACTAGGTCCGCAAGTCTGTTATGTGCAAGTTCCCAAGAGTTTGTATGGGA 1860
Db 1801 TGGATCAAACTAGGTCCGCAAGTCTGTTATGTGCAAGTTTCCCAAGAGTTTGTATGGGA 1860

QY 1861 TTGATAGGAATGATTCGATATGCAAAAGCAAGCACTGTCTTTTGTATTAATTTAGGG 1920
Db 1861 TTGATAGGAATGATTCGATATGCAAAAGCAAGCACTGTCTTTTGTATTAATTTAGGG 1920

QY 1921 GCCTTGACGGCATTCAGAGGACCGATTATGTGGAACTGGTTGTTTCTCAACAGACAG 1980

Db	1921	GCCTTGACGGCATTCAAGGACCACTTATGTGGAACATGGTGTGTGTTTCAACGAACAG	1980
Qy	1981	CTATCTATGGTTATGAGCCCCCAATTAAGGCGAAGAACCCAGGTTTCTTGGCATCACTAT	2040
Db	1981	CTATCTATGGTTATGAGCCCCCAATTAAGGCGAAGAACCCAGGTTTCTTGGCATCACTAT	2040
Qy	2041	GTGGGGGCAAGAAAGCAAGCAAGTCAAAAGAAAAGAGCTCAGATAAGAAAAGTCGA	2100
Db	2041	GTGGGGGCAAGAAAGCAAGCAAGTCAAAAGAAAAGAGCTCAGATAAGAAAAGTCGA	2100
Qy	2101	ACRAGCATGGGACAGTCTGTTCAGTATTCAATCTCGAAGACATAGAGAGGGTGTG	2160
Db	2101	ACAAGCATGGGACAGTCTGTTCAGTATTCAATCTCGAAGACATAGAGAGGGTGTG	2160
Qy	2161	AAGGTGCTGGGTTTGATGATGAGAAAATCAGTTCATGTCTCAATCAGCTTAGAGAAGA	2220
Db	2161	AAGGTGCTGGGTTTGATGATGAGAAAATCAGTTCATGTCTCAATCAGCTTAGAGAAGA	2220
Qy	2221	GATTTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGTTCCTC	2280
Db	2221	GATTTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGTTCCTC	2280
Qy	2281	AGTCCTCCACTCCAGATCTCTTTTGAAGNAGCTATCCATGTCATTAAGTTTGGCTATG	2340
Db	2281	AGTCCTCCACTCCAGATCTCTTTTGAAGNAGCTATCCATGTCATTAAGTTTGGCTATG	2340
Qy	2341	AGGACAAAGTCTGAATGGGGAACCTGAGATTTGGTTCGATCTTGGATCTGTCACAGAAGATA	2400
Db	2341	AGGACAAAGTCTGAATGGGGAACCTGAGATTTGGTTCGATCTTGGATCTGTCACAGAAGATA	2400
Qy	2401	TTCTTACTGGAATCAAGATGCAAGCAGAGGCTGGCGTTCACTCTATTGCATGCCCAAGC	2460
Db	2401	TTCTTACTGGAATCAAGATGCAAGCAGAGGCTGGCGTTCACTCTATTGCATGCCCAAGC	2460
Qy	2461	GCCCAGCTTTCAAGGGATCTGGCCCCCATCAATCTTTTCAGATCGTCTGAAACAAGTGTCTC	2520
Db	2461	GCCCAGCTTTCAAGGGATCTGGCCCCCATCAATCTTTTCAGATCGTCTGAAACAAGTGTCTC	2520
Qy	2521	GGTGGGCTCTCGGTTCTGTTGMAATCTTTTCAGCGGGCATGCGCCCTTATGATGCTGGCT	2580
Db	2521	GGTGGGCTCTCGGTTCTGTTGMAATCTTTTCAGCGGGCATGCGCCCTTATGATGCTGGCT	2580
Qy	2581	ACGAGAGGGCGCTCAAGTTCCTGGAGAGATTCGTTACATCAACACACCATTTACCCAC	2640
Db	2581	ACGAGAGGGCGCTCAAGTTCCTGGAGAGATTCGTTACATCAACACACCATTTACCCAC	2640
Qy	2641	TAACCTCTCCCGCTCTAGTCTATTGTATATTGCTGCTATCTGCTCACTGGA	2700
Db	2641	TAACCTCTCTCCCGCTCTAGTCTATTGTATATTGCTGCTATCTGCTCACTGGA	2700
Qy	2701	AGTTCAATCTGCAGAGATTAGCAACTTGGCCAGTATCTGGTTCATTGCTCTCTCTCTCT	2760
Db	2701	AGTTCAATCTGCAGAGATTAGCAACTTGGCCAGTATCTGGTTCATTGCTCTCTCTCTCT	2760
Qy	2761	CAATTTTCGCCACTGGTATCCTTGGATGAGGTGGAGTGGTGTGGCATTTGACGAGTGGT	2820
Db	2761	CAATTTTCGCCACTGGTATCCTTGGATGAGGTGGAGTGGTGTGGCATTTGACGAGTGGT	2820
Qy	2821	GGAGGAATGAACAGTTCTGGGTCAATTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTC	2880
Db	2821	GGAGGAATGAACAGTTCTGGGTCAATTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTC	2880
Qy	2881	AGGTCCTCTGAAAGTCTTGGCGGTATCGACACCACTTCACTGTCACTCAAGGCTA	2940
Db	2881	AGGTCCTCTGAAAGTCTTGGCGGTATCGACACCACTTCACTGTCACTCAAGGCTA	2940
Qy	2941	ATGACGAAGAGGCGACTTTTGTCTGAGCTCTACATGTTCAAGTGGAGAGGCTTCTCATCC	3000
Db	2941	ATGACGAAGAGGCGACTTTTGTCTGAGCTCTACATGTTCAAGTGGAGAGGCTTCTCATCC	3000
Qy	3001	CTCCGACGACATTTTGATCAATTAACATGGTGGTGTGCTGGCACCTCTCAGCCA	3060

Db	3001	CTCGAGCAGCACAATTTTGATCATTAACATGGTTGGTGTCTGTTGGCAGCACTCTTACGCCA	3060
Qy	3061	TCAACAGTGGTTACCAATCATGGGGGGCGCTCTTTGGGAAGCTCTTCTTTGCCTTCTGGG	3120
Db	3061	TCAACAGTGGTTACCAATCATGGGGGGCGCTCTTTGGGAAGCTCTTCTTTGCCTTCTGGG	3120
Qy	3121	TGATTGTCACTTATACCCATTCTCAAGGTCCTTATGGGCAGGCAAAACCGCACACCGA	3180
Db	3121	TGATTGTCACTTATACCCATTCTCAAGGTCCTTATGGGCAGGCAAAACCGCACACCGA	3180
Qy	3181	CGATTGTCATCGTCTGGGCTGCTCCTCGTCTCTATCTTCTCCTTGCCTGTGGTTCGTG	3240
Db	3181	CGATTGTCATCGTCTGGGCTGCTCCTCGTCTCTATCTTCTCCTTGCCTGTGGTTCGTG	3240
Qy	3241	TTGATCCAATTCACCTACCCGCTCTCGCTGGGCCAAATATCCAAACCTGTGGGCATCAACTGCT	3300
Db	3241	TTGATCCAATTCACCTACCCGCTCTCGCTGGGCCAAATATCCAAACCTGTGGGCATCAACTGCT	3300
Qy	3301	AGGAAAGTGGGAGTTTGTAGAGACAGAAAATATAACAGTGATCGAGCAACAACCCGGCGGA	3360
Db	3301	AGGAAAGTGGGAGTTTGTAGAGACAGAAAATATAACAGTGATCGAGCAACAACCCGGCGGA	3360
Qy	3361	GCCAGAGAAATATTTATCTTCGGGTTGTGAATTACTACGTTTGGAGAAAGTTGTCAAAAATG	3420
Db	3361	GCCAGAGAAATATTTATCTTCGGGTTGTGAATTACTACGTTTGGAGAAAGTTGTCAAAAATG	3420
Qy	3421	AGAAAACACATTTGTAAATAGATGTAATAGACCATCTACCGTTTTCATGAGGTTAAAGCTC	3480
Db	3421	AGAAAACACATTTGTAAATAGATGTAATAGACCATCTACCGTTTTCATGAGGTTAAAGCTC	3480
Qy	3481	TTCTTTTTTTGGACAAAGAACTCATTTGGTAAACCTATAGGAAATTTTCCATGAGGCA	3540
Db	3481	TTCTTTTTTTGGACAAAGAACTCATTTGGTAAACCTATAGGAAATTTTCCATGAGGCA	3540
Qy	3541	CTTTGGATTGTAGGAATGGACCTATGAAATGTTGTATTATTTATTTATATAAAATTATTC	3600
Db	3541	CTTTGGATTGTAGGAATGGACCTATGAAATGTTGTATTATTTATTTATATAAAATTATTC	3600
Qy	3601	TGTCCTTCACATTTTGGAGGAGTTTT	3626
Db	3601	TGTCCTTCACATTTTGGAGGAGTTTT	3626

QY	1135	TCAGTACAGTGCAGCCCTTTGAAGGAGCCACCTATCGTCACTGCACACATGCTGCTATCCA	1194
Db	1512	TCAGTACAGTGCAGCCCATGAAGGACCTCTCTCTGTTACTGCCAATACCGTGTGTCCA	1571
QY	1195	TTCTTGCTGTTGATATCCCGTGGACAAGGTCTCTTGCTATGTATCTGATGACGAGCTT	1254
Db	1572	TTCTTGCTGTTGATATCCCGTGGACAAGGTCTCTTGCTATGTATCTGATGACGATGGTCTG	1631
QY	1255	CAATGCTGACTTTTACGCAATTTGGCTGAGACTTCAGAGTTTCTAGGAATAGGTACCAT	1314
Db	1632	CAATGCTGACTTTTATGCAATTTGGCTGAGACTTCAGAGTTTCTAGGAATAGGTACCAT	1691
QY	1315	TTGTGAAGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTCTTTTCCAGAAAATTTG	1374
Db	1692	TCGTTAAGAAGTATTAACATTTGAGCCAGAGCTCTCGAGTGTACTTCTCCCAAGAAATCG	1751
QY	1375	ATTACTGAAAGACAAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGCGCATGAAGAGAG	1434
Db	1752	ATTACTGAAAGACAAAAGTCCAGCTTCAATTTGTTAAAGACCGCTCGTGCATGAAGAGAG	1811
QY	1435	RATATGAGAATTTAAATCAGATTAATGCCCTAGTTTCTTAAGGATTTGAAGTCCCGG	1494
Db	1812	AATATGAAGAATTTCAAAGTTAGGATTAATGGCCCTTGTGCTAAGGCACAGAAAGTCCCGG	1871
QY	1495	AGGAAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAACAATACCGGGATCATC	1554
Db	1872	AGGAGGGATGGATCATGCAAGATGGCACACCATGGCCAGGAACAATACCGGGATCATC	1931
QY	1555	CTGGAATGATTGAGTTTCTTGCTGCTACAGTGTGGCTTGATACTAGGGGTAATGAGC	1614
Db	1932	CTGGAATGATTGAGTTTCTTGCTGCTACAGTGTGGCTTGATACTAGGGGTAATGAGC	1991
QY	1615	TCCCGGTTTATGTTGCTCTGTAAGAGCTCTGGGTTCAGACACCAAGAAGG	1674
Db	1992	TTCCCGGCTGGTCTACGATATCTGTAAGAGCTCTGGGTTCAGACACCAAGAAGG	2051
QY	1675	CTGGTGCCATGAATGCCCTTGTCTGCTGCTCAGCTGTCTTACTAAATGAGCAATACATGT	1734
Db	2052	CTGGTGCCATGAATGCTCTTGTCTGCTCAGCTGTCTTACTAAATGAGCAATACATGT	2111
QY	1735	TGAATCTTGATGTTGATCACTACATCAACACAGCAAGGCTGTCCGAGAAGCTATGTCT	1794
Db	2112	TGAATCTTGATGTTGATCACTACATCAACACAGCAAGGCTGTCCGAGAAGCTATGTCT	2171
QY	1795	TCCTAATCGATCCAAACCTAGTCCGAGAGCTGTATGTGAGTTTCCCAAGAAGTTTG	1854
Db	2172	TCCCTATGGATCCAAACCTAGGAAGGAGTGTCTGTTATGTTCAAGTCCCAAGAGTTTCG	2231
QY	1855	ATGGGATTTGATAGGAATGATCGATATGCAACAGGAACAACCTGTCTTTTGTATATAACT	1914
Db	2232	ATGGATTTGATAGGAATGATCGATATGCAACAGGAACAACCTGTCTTTTGTATATAACT	2291
QY	1915	TGAGGGGCTTTGACGGCATTTCAAGGACAGTTTATGTCGGAACCTGTTTGTATTTCAACA	1974
Db	2292	TGAGGGGCTTTGATGGCATCAAGGACAGTTTATGTCGGAACCTGTTTGTATTTCAACA	2351
QY	1975	GAAAGCTTATCTATGTTTATGAGCCCAATTA --- AGCCGAAGAACCGAGTTTCTTGG	2031
Db	2352	GAACTGCTTATGTTTATGAAACCCCAATTAAGCAAGAAAGAGGAGGAGTTTCTTGT	2411
QY	2032	CATCACTATGTTGGGGCAAGGAAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAGA	2091
Db	2412	CATCACTATGTTGGGGCAAGGAAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAGA	2471
QY	2092	AAAAGTCGAACAGCAATGAGCAGTTCTGTTTCCAGTATTTCAATCTCGAAGACATAGAGG	2151
Db	2472	AGAGTCAAAACAGCAATGAGCAGTTCTGTTTCCAGTATTTCAATCTCGAATATAGAGG	2531
QY	2152	AGGGTGTGAAGGTGCTGGGTTTGTATGATGAAATCAGTTCTCATGTCTCAATAGAGCT	2211
Db	2532	AGGGTGTGAAGGTGCTGGATTCGATGATGAAATCAGTTCTCATGTCTCAATAGAGCT	2591

QY	2212	TAGAGAAGAGATTTGGCCAGTCCAGAGCATTTGTTGCCCTCACTCTGATGGAATATGGTG	2271
Db	2592	TGAGAAGAGATTTGGCCAGTCTGAGCGGTTGTTGGCTCCACTCTCATGGAATATGGTG	2651
QY	2272	GTGTTCTCCTCAGTCCCTCCATCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCATAAGTT	2331
Db	2652	GTGTTCTCCTCAATCTGCAACCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCATAAGTT	2711
QY	2332	GTGGCTATGAGGAACAAGTCTGAATGGGAACTGAATTTGGTGGATCTATGATCTGTCA	2391
Db	2712	GTGGCTATGAGGAACAAGTCTGAATGGGAACTGAATTTGGTGGATCTACGGTCCGTGA	2771
QY	2392	CAGAAAGATATCTTACTGSAATCAAGATCGACGCAAGAGGCTGGGTTTCACTCTATTGCA	2451
Db	2772	CAGAAAGATATCTTACTGSAATCAAGATCGACGCGAGGCTGGAGATCAATCTACTGCA	2831
QY	2452	TGCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATCTCTGAACC	2511
Db	2832	TGCCCAAGCGCCAGCTTTCAAGGGATCTGCCCTTCAATCTTTTCAGATCTCTTAACC	2891
QY	2512	AAGTCTCGGTGGGCTCTCGGTCTCTGTGAAATTTCTTTTCAGCCGGCATTTGCCCTTAT	2571
Db	2892	AAGTCTCTCGGTGGGCACTTGGTTCTGTGAAATTTCTTTTCAGTCCCATTTGCCCATAT	2951
QY	2572	GGTATGGCTACGGGGCGCCTCAAGTTCTCTGGAGAGATTCGCTTACATCAACACCA	2631
Db	2952	GGTATGGCTATGGAGGCGCCTTAAGTTCTTTGGAGAGATTTGCCCTACATCAACACCA	3011
QY	2632	TTTACCCACTAACCCTCTCTCCGCTTCTAGTCTATTTATTTATTTGCTGCTATCTCTGCG	2691
Db	3012	TTTATCCATTCGATTCGATCCGCTTCTCATATACCTGTGTTTGGCTGCTATCTGTTGC	3071
QY	2692	TCATGGAAGTTTCATGATGCGAGAGATTAGCAACTTGGCCAGATCTGTTTCATTCGCG	2751
Db	3072	TCATGGAAGTTTCATTCATCCAGAGATTAGCAACTTGTAGTATTTGGTTTCATCTCTC	3131
QY	2752	TCTTCTTTTCAATTTTCGCCACTGATCTCTTCTGATGAGTGGGAGTGGTGTGCAATG	2811
Db	3132	TCTTCAATTTTCAATTTTTCGCCACTGATCTCTTCTGATGAGTGGGAGTGGTGTGCAATG	3191
QY	2812	ACGAGTGTGGAGGAATCAACAGATTTCTGGGCTCATTTGGAGGTATCTCTGCACATCTGTTG	2871
Db	3192	ATGAGTGTGGAGGAATCAACAGATTTCTGGGTTATTTGGAGGTATATCTCGGCATCTTTTG	3251
QY	2872	CCGTCTTTTCAAGGTTCTCTGAAGGTTCTGCGGTATCGACACCAACTTCACTGCACT	2931
Db	3252	CCGTCTTTTCAAGGTTCTCTGAAGGTTCTGCGGTATCGACACCAACTTCACTGCACT	3311
QY	2932	CAAGGCTAATGACGAAGAGCGACTTTGCTGAGCTCTACATGTTTCAAGTGGAGCAAGC	2991
Db	3312	CAAGGCTTCTGATGAGATGGGACTTTGCTGAGCTCTACATGTTTCAAGTGGAGCAAGC	3371
QY	2992	TTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGTTGGTGTCTGCTGCGCACT	3051
Db	3372	TTCTCATCCCAAGGACCAATCTTGATCATTTACCTGCTGCGTGTGTTGCTGCTGCT	3431
QY	3052	CCTAGGCTTCAACAGTGGTTTACCAATCAAGGGGCGCTCTTTGGGAAGCTCTTCTTTG	3111
Db	3432	CATAGGCTTCAACAGGCGCTTACCAATCAATGAGGCGCTCTTTGGGAAGCTCTTCTTTG	3491
QY	3112	CTTCTGGGTTGTTTCACTTATACCCATTCCTCAAGGCTCTTATGGGAGGCAAAACC	3171
Db	3492	CTTCTGGGTTGTTTCACTTATACCCATTCCTCAAGGCTCTTATGGGAGGCAAAACC	3551
QY	3172	GCACACGAGATTTGTCATCTGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3231
Db	3552	GCATCCGACCATCTGTTGTTGTTGGGCAATCTCTCTGCTCTCTCTCTCTCTCTCTCTCT	3611
QY	3232	GGGTTCTGTTGATTCATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3291
Db	3612	GGGTTCTGATTCATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3671
QY	3292	TCAACTGCTAGGAAGTGGGAGTTTGTAGAGACAGAAAATATAACAGT	3339

Db	2761	GGTGTGGCATCGATGAGTGGTGAGGAATGAACAGTCTCTGGGTATTGGAGGTATATCT	2820
Qy	2859	GCATCTGTTTGGCCGCTTTCAGGGTCTTCTGAAGTGCTTCCGGTATCGACACCAAC	2918
Db	2821	CGCATCTTTTGGCCGCTTCCAGGGTCTCTCTCAAGGTGCTTGTGTATCGACACCAAT	2880
Qy	2919	TTCACTGTCACTCAAAAGGCTTAATGACGAAGAAGCGCACTTGTCTGAGCTCTACATGTC	2978
Db	2881	TTCACTGTCACTCAAAAGGCTTCTGATGAAGATGGCGACTTGTCTGAGCTCTACATGTC	2940
Qy	2979	AAGTGGACGAGCTTCTCATCCCTCCGACGACCAATTTTGATCAATAACATGGTGTGTC	3038
Db	2941	AAGTGGACAAAGCTTCTCATCCCAACGACCAATCTTGATCATTAACCTGGTGGTGT	3000
Qy	3039	GTTGCTGGCACTCTCAAGCCATCAACAGTGTATACCAATCATGGGGCGCTCTTTGGG	3098
Db	3001	GTTGCTGGTATCTCATAGCGATCAACAGCGGTACCAATCATGGGGACCGCTCTTTGGC	3060
Qy	3099	AAGCTCTCTTTTGGCTTCTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGGTCTTATG	3158
Db	3061	AAGCTCTCTTTTGGCTTCTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGGTCTTATG	3120
Qy	3159	GGCAGGCAAAACGACACGACGATTTGTCATCGTCTGGGTGTCCTCCCTGCTTCTATC	3218
Db	3121	GGTCGGCAAAACGCACTCCGACCATCGTTGTTGTTGGGCAATCTTCTGGCTTCGATC	3180
Qy	3219	TTCTCCCTTGCTGGGGTTCGTGTGATCCATTCACTACCGCTCTCGTGGCCCAATATC	3278
Db	3181	TTCTCATTTGCTGGGGTTCGCATCGATCCATTCAACACCGGTGTCACCGCCAGATACC	3240
Qy	3279	CAAACTGTGGCATCAACTGCTA	3301
Db	3241	CAAACTGTGGCATCAACTGCTA	3263
RESULT 4			
AF200533			
LOCUS	AF200533	3795 bp	linear PLN 31-AUG-2000
DEFINITION	Zea mays cellulose synthase-9 (Cesa-9) mRNA, complete cds.		
ACCESSION	AF200533		
VERSION	AF200533.1	GI:9622889	
KEYWORDS			
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 3795) Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S., Xoconostle-Cazares,B. and Delmer,D.P.		
TITLE	A comparative analysis of the plant cellulose synthase (Cesa) gene family		
JOURNAL	Plant Physiol.	123 (4),	1313-1324 (2000)
MEDLINE	20398328		
PUBMED	10938350		
REFERENCE	2 (bases 1 to 3795) Dhugga,K.S. and Helentjaris,T.G.		
AUTHORS	Direct Submission		
TITLE	Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1. 3795 /organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"		
gene	1. .3795 /gene="Cesa-9"		
CDS	238..3477 /gene="Cesa-9" /note="Cesa-9"		

		/codon_start=1		/product="cellulose synthase-9"			
		/protein_id="AAF89969.1"					
		/db_xref="GI:9622890"					
		/translation="MEGDADGVKSRGGQVQCIGDGVGTAAEGDVFTACDVCGGP					
		VCRPCVYERKDGTOACPOCKNKYKRHKGPSAIRREGDDTDADDADEFNYPSAGNDD					
		OKOKIADRMRSMNAGSDYRKPYSDEIGLTKYDSGEIPRGVIPSVTNMQISGE					
		IFGASPDHMSFTGNIGRRAPPYNNHSSNRPSPSGSVGNVANKRVDGWMKQDK					
		GITPMNTGTSIPSEGRGVGDIDASDIDNNEADALNDETQPLSRKVPPLPSRINPYR					
		MYIVLRILVLSIFLHVRIITPNVRYNAYPLWLLSVICBIWFALSWILDQFPKWFPIKNET					
		YLDRLALRYDRGEPSQALAVDIFVSTDVPMKEPPLVANTLILSLAVDIPDKVSCY					
		VSDGAAMLTDFALAEISFARKWPVFKKYNIEIPRAPEWFSQKIDYLDKQVHFSFV					
		KDRAMRVEEFKIRVINGLVAKAQVPEGIMODGTPWGNTRDHPQMLOVFLGH					
		SGGLDEGNEIPLRVYVSRKRPFGQHKKACAMNALVRSVAVLTNGQYMLNLDCHY					
		INNSKALREAMCLMDPNLGRSVCYVQFQRFQDIDNRNRYANRNTVFDINLRGLDG					
		IDGPVVTGCVFNRTALYGEYPIIKQKGGFLLSLCCGRKKSGKSGDKSKQKH					
		VDSVPVFNLEDIIEEGVAGGDDXSLMLMSLEKRFQCSAAVFASITMEYGGVQF					
		SATPELLEAETHVISCYEDKIEMGTEIGWYGSVEDILITGFQKHARGWRSIYCMP					
		KRPAPKGSAPINLSRLNOVLWALGSLVILPSRHCPVWYGVGRKLFERFAYINTT					
		IYPLTSLPILLYCILPAICLLTGKFIIPISINPASIFWISLFSIFATGILEMRWNGV					
		GIDWRNRQFWVIGISAHLPFAVQGLLKLVLAGIDTNETVTSKASDEGDEPAELWYF					
		KMTLLIPTTILILINLVGVAGISYAINSGYQSWGLPFGKLFFAFWVIIVHLVPLFKG					
		LMGRNRTPTIVVWAILLASIFSLWVRIDPFTNRTVTPGDTNRTGGINC"					
		ORIGIN					
		Query Match 71.6%; Score 2595; DB 8; Length 3795;					
		Best Local Similarity 87.1%; Pred. No. 0;					
		Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;					
Qy	36	CGGGTGGATCGCTCTGCCGCGCATGGACGGCGACGGCGACGCCCTTGAAAGTCGGGAGG	95	696	AAAAAGACGACGACGAAGGGTGCATTTCCATGACTAATGGGACAAAGCATTTGCTCCCTCT	755	
Db	214	CGGTTCTGGCGAGCTCGCTGCCATGGAGGGCGACGGCGAGTGAAGTCGGGAGG	273	874	AAAAAGACGACGACGAAGGGAAACAATTTCCATGACGAATGGCAAGCATTTGCTCCCTCT	933	
Qy	96	CACGGGCGGGGACGTTGCCAGATCTGCGCGACGCGCCCTGGGCACCAAGTTGGACGGC	155	756	GAAGTTCGGGCGCTTACTGACATCGATGCAATCTACTGATTAACAATGAAGAAGCGCTTTA	815	
Db	274	CGCGGGGAGGCGAGGTGCCAGATCTGCGCGGATGGCGTGGGCACCTACGGGGAGGGA	333	934	GAGGGCCGGGGTGTGTGTATTTGATGCAATCAACTGATTACAACATGAAGATGCTTTA	993	
Qy	156	GAGCTTTACCGCTCGGACGCTGCGCTTCCCGTCTGCGCCCTGCTACGAGCAC	215	816	CTGAATGATGAACACTCGCCAGCTCTATCTAGAAAAAGTCCCATCTCTCTCCCAAAATA	875	
Db	334	GAGCTTTACCGCTCGGACGCTGCGGGTGTCCCGTGTGCGCCCTGCTACGATAC	393	994	TTAAAGATGAACACTCGCCAGCTCTATCTAGAAAAAGTCCCATCTCTCTCCCAAGATA	1053	
Qy	216	GAGCGCAAGGAGGACCCAGGCGTCTCCAGTGAAGACCAAGTACAAGCGCCACAGA	275	876	AATCCCTACAGAAATGCTGCTTCTGCGGTTGGTTTCTTAAGACATTTCTTCGCACTAC	935	
Db	394	GAGCGCAAGGACGCGACACACAGCGTGCCTCCAGTGAAGACCAAGTACAAGCGCCACAG	453	1054	AATCCATACAGGATGCTGCTTGTGCTACGATTGTTTCTTAAGCATCTTCTTTCGCACTAC	1113	
Qy	276	GGAGGCCACGATCCCGGGGAGGAGGCGACACTGATGCCGATGATGGTAGTAC	335	936	CGTCTCACAAATCTGTGGTGAATGATACCACTGTGGCTTTTATCTGTATATGTGAG	995	
Db	454	GGGAGTCCAGCGATCCGAGGGAGGAAAGGAGCGATACTGATGCCGATGATGTCGCGAC	513	1114	CGGATCACAAATCTGTGGTGAATGATACCACTGTGGCTTTCTATCTGTATATGTGAG	1173	
Qy	336	TTCAACTACCTGCTATCGACCTGAGGACCAAGACAGAGATGCTGACAGGATCGCG	395	996	ATTGTGTTGCTTTATCTCTGGATACGATCAGTTCCCGAAGTGGTTTCCAATCAACCGG	1055	
Db	514	TTCAACTACCTGCTATCGCAATGACGACACAGNAGCAGAGATGCTGACAGATCGCG	573	1174	ATCTGTTGCTCTTTCTCGGATTTGGATCAGTTTCCAAAGTGGTTTCCAATCAACCGC	1233	
Qy	396	AGCTGGCGCATGAACACCGGGGGCAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGC	455	1056	GAGACCTACTCTTGATAGACTGGCTTTAAGGTATGACCGGAGAGGTGAACCGTCTCAGTTG	1115	
Db	574	AGCTGGCGCATGAATGCTGGGGCGAGCGGGATGTTGGCGCCCAAGTATGACAGTGGT	633	1234	GAGACTTACTCTTGATAGACTCGCATTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTG	1293	
Qy	456	GAGATCGGCTCTCAGATATGACAGTGGAGAGATCCCTAGGGGATACGCTCCCTTACGTC	515	1116	GCTGCTGTTGACATATTTGTCAGTACAGTCGACCCCTTGAAGGAGCACCATTATCGTCACT	1175	
Db	634	GAGATCGGCTTACCAAGTACGAGTGGTGGATCCCTCGGGGATACATCCCGTCAGTC	693	1294	GCTGCTGTTGACATTTTGTGCTGCTGCGACCCCAATGAAGGAGGCTCCTCTGTGCACT	1353	
Qy	516	ACCAACAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCCGCTGATCATCATGATGCTCC	575	1176	GCCAAACACTGTGCTATCTGCTGTTGATTAATCCCGTGGACAAAGTCTCTTGTCTAT	1235	
Db	694	ACTAACAGCCAGATTTCCGAGAAATCCCTGGTGTCTCCCTGACCATCATATGATGCT	753	1354	GCCAAACACTGTGCTATCTGCTGTTGGAATCTCTGCTGGAATCTCTGATGATGAGTCTCTTGTCTAT	1413	
Qy	576	CCTACGGGGACATCAGCAGACGCTGCTCCGTTTCCCTATGTGAATCAATCAACAAATCCG	635	1236	GTATCTGATGACGAGGCTTCAATGCTGACTTTTTCGCACTTTGACGCAATGCGTGAAGTCTT	1295	
Db	754	CCTACTGGGAACATTTGGCAGCGCGCCCAATTTCCCTATATGAATCATCAATCAATCCG	813	1414	GTATCTGATGATGAGCTGCTATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTT	1473	
Qy	636	TCAGGAGATTCTCCGAGTATTTGGGAATGTTCCCTGGAAAGAGAGATTGATGGCTGG	695	1296	GCTAGGAATGGGTACATTTGTGAAGAGATGACATGAACCCGAGAGCTCCGAGTTT	1355	
Db	814	TCGAGGGAATTTCTGTGTAGCGTTGGGAATGTTCCCTGGAAAGAGAGAGGTTGATGGCTGG	873	1474	GCTAGGAATGGGTGCTCATTTGTTAAGAGTACAACATTAACCTAGAGCTCCTGGAATGG	1533	

KRPAFKGSAPINLSDRLNQVLRWALGWSVEILFSRHCPLWGYGGRUKFLERFAYINVT
IYPLTSPLLIYCIIPALICLLTKGKIIPEISNFASIWFTSLFISIPATGILEMRWGV
KIDWWRNEQWVIGGISAHLFAVQGLLKVLGADITNFTVTSKASDGDGFAELXYMF
GWTTLIPPTTILINLVGVVAGISYAINSGVSGWCPFGKLFAPFWIWLHYPLPKG
IMGKRNPTTIVVWAILLAFSLIWLVRIDPFTNRVTGDPFTTTCGINC"

ORIGIN

Query Match 71.6%; Score 2595; DB 6; Length 3799;

Query Macchi	71.0%,	Score 4333,
Best Local Similarity	87.1%	Pred. No. 0:

BEST LOCAL SIMILARITY	87.1%	FREQ NO: 0,
Matches 2871:	Conservative	0: Mismatches 420: Indels 4: Gaps 2:

36	CGGGTTGGATCGCTCTGCGCGCCATGACACGGCAGCGGACGCGCTGGAAGTCCGGCAGG	95
214	CGTTCTGGCGAGCTCGCCTGCATGAGGGCGACGGCGACGGGTGAAGTCGGGGAGG	273
96	CACGGGGCCGGGGAAGTGTGCGCGATCTTGCGCCGACGCCCTGGGGACACACGTTGGACGGC	155
274	CGCGGGGAGGGCAGGTGTGCCAGATCTCGGCGCATGGCGTGGGCACCTACGCGCGAGGGGA	333
156	GACGCTTTCACCGCCTGCGACGCTCTGCGCGCTTCCCGGTCTCCGCGCTCTCGCGCCCTGTACGAGCAC	215
334	GACGCTTTCACCGCCTGCGACGCTCTGCGGGTCTCCCGGTGTGCGCGCCCTGCTACGAGTAC	393
216	GAGCGCAAGGAGGGACCCACGGCCTGCCTCCAGTCCAAAGACCAAGTACAAGCGCCACAGA	275
394	GAGCGCAAGGACGGCACACACGGGTGCCCCAGTGCATAAACAAGTACAAGGGCCACAAG	453
276	GGGAGCCGACGATCCGCGGGGAGGAAGCGGACGACACTGATGCCGATGATGGTAGTGAC	335
454	GGGAGTCCAGCGATCCGAGGGGAGGAAGGAGACGATACTGTATGCCGATGATGCTTAGCGAC	513
336	TTTCAACTACCTTCATCTCGCACTGAGGACACGAGACGAGAATGCTGCACAGATGCGC	395
514	TTCAACTACCTTCGATCTGGCAATGACGACACGAGACGAGAATGCTGCACAGATGCGC	573
396	AGCTGGCGATGAACCCGGGGCAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGC	455
574	AGCTGGGCGATGAATGCTGGGGGACGCGGGGATGTTGCCCGCCCCAAGTATGACAGTGGT	633
456	GAGATCGGCTCTCCAAGTATGACAGTGAGAGATCCCTAGGGGATAGCTCCCTTCAGTC	515
634	GAGATCGGGCTTACCAAGTACGACAGTGGTGAGATCCCTCGGGGATACATCCGCTCAGTC	693
516	ACCAACAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCACATGATGCC	575
694	ACTAACAGCCAGATTCGCGAGAAATCCCTGGTGCTTCCCTTGACCATCATGATGHT	753
576	CCTACGGGGAACTCAGCACGCTGCTCGGTTTCCCTATGTGAATCATTTCCACCAATCCG	635
754	CCTACTGGGAACTTGGCAGCGCGGCCCATTTCCCTATATGAATCATTCATCAATCCG	813
636	TCAAGGAGTTCCTCGGCAGTATTTGGATTTGGTTCCTGGAAAGAGAGAGTGTATGGCTGG	695
814	TCGAGGGAAATTCCTGGTAGCTTTGGGAAATGTTGCTGGAAAGAGAGGGTGTATGGCTGG	873
696	AAATGAGCAGGACAGGGTGCATTTCCATGACTTAATGGGACAAAGCATTCCTCCCTCT	755
874	AAATGAGCAGGACAAAGGAAACAAATTCCTGACGAAATGGCAAGCATTCCTCCCTCT	933
756	GAAGGTCGGGCAGCTACTGACATCGATGCTACTGAATACAACATGGAAGACGCTTTA	815
934	GAGGCGGGGTGTTGGTGATATTGATGCATCAACTGATTACAACATGGAAGATGCCTTA	993
816	CTGAATGATGAACCTCCGAGCCTCTACTAGAAAAGTCCCATTTGCTTCCTCCAAATA	875
994	TTAACGATGAACCTCCGAGCCTCTACTAGAAAAGTCCCATTCCTTCCTCCAGGATA	1053
876	AATCCCTACAGATGGTCAATTGCTCGGTTGGTTGTTCTAAGCATCTTCCTGCACTAC	935
1054	AATCCATACAGGATGGTCATTTGCTACGATTGATTGTTCTAAGCATCTTTGCACTAC	1113
936	CGTCTCAAAATCCTGTGCGTAATGCATACCACTGTGGCTTTTATCTTTTATGTGAG	995

```
QY 2076 AGGAGTCAATAGAAAAAGTCCAAACAAGCATGTGGACAGTTCCTGTTCCAGTATTCAT 2135
Db |||||
QY 2253 --GGGCTCAGACAAGAAAAAGTCACAGAACAATGTGACAGTTCCTGTCGCCAGTATTCAT 2310
Db |||||
QY 2136 CTCGAAGACATAGAGAGGGTGTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTC 2195
Db |||||
QY 2311 CTTGAAGATATAGAGGAGGAGTTGAAGGCGCTGGATTTGATGATGAGAAATCATTCTT 2370
QY 2196 ATGCTCTCAATGAGCTTAGAGAAGAGATTTGGCCAGTCACAGCATTTGTTGCTCCACT 2255
Db |||||
QY 2371 ATGCTCTCAATGAGCTTAGAGAAGAGATTTGGCCAAATCTCGAGCTTTTGTGGCTCCACT 2430
QY 2256 CTGATGGAATPANGGTGGTCTCTAGTCTCCACTCCAGAAATCTCTTTGAAAGAAGCT 2315
Db |||||
QY 2431 CTGATGGAATATGGTGGTCTCTAGTCTGCGCACTCCAGAAATCTCTCTGAAAGAAGCT 2490
QY 2316 ATCCATGTCATAGTTGTGGCTATGAGGACAAAGCTGAATGGGAACTAGAGATTGGTTGG 2375
Db |||||
QY 2491 ATCCATGTCATAGTTGTGGCTACGAGGACAAAGATTGAATGGGAACTAGAGATTGGTTGG 2550
QY 2376 ATCTATGGATCTGTACAGAAGATATCTTACTCGAATCAAGATGACGACGAGAGCTGG 2435
Db |||||
QY 2551 ATCTATGGTCTGTGACGAAGATATCTACTGGGTTCAAGATGACGACGAGAGCTGG 2610
QY 2436 CGTTCAGTCTATTGCAATGCCCAAGCGCCAGCTTTCAAGGATCTCGCCCATCAATCTT 2495
Db |||||
QY 2611 CGGTGATCTACTGATGCTTAAGCGCGCGCTTCAAGGATCGCTCCATCAATCTC 2670
QY 2496 TCAGATGCTGTGAACCAAGTCTGCGTGGGCTCTCGGTTCTGTGAATCTTTTCAGC 2555
Db |||||
QY 2671 TCAGACCGTCTGAACCAAGTCTGCGTGGGCTCTCGGTTCAAGTGAATCTTTTCAGC 2730
QY 2556 CGGCATTGCCCTTATGATGATGCTACGAGGCGCTCAAGTCTCTGAGAGATTCGCT 2615
Db |||||
QY 2731 CGGCATTGCCCTTATGATGATGCTACGAGGCGCTCAAGTCTCTGAGAGATTCGCT 2790
QY 2616 TACATCAACACACCACTTATACCACTAATCTCTCCCGCTCTAGTCTATTGTTATATG 2675
Db |||||
QY 2791 TACATCAACACCACTTATACCACTAATCTCTCCCGCTCTAGTCTATTGTTATATG 2850
QY 2676 CTTGCTATCTGCTGCTCACTGGAAGTTCATCATGCCAGAGATTAGCACTTGGCCAGT 2735
Db |||||
QY 2851 CTTGCTATCTGCTGCTCACTGGAAGTTCATCATGCCAGAGATTAGCACTTGGCTAGT 2910
QY 2736 ATCTGTTTCATTGGCTCTCTCTTCAATTTTCGCCACTGGTATCTCTTGAGATGAGTGG 2795
Db |||||
QY 2911 ATCTGTTTCATTCTCTCTTCACTCTGATCTTTCGCCACGGGTATCTCTGAGATGAGTGG 2970
QY 2796 AGTGGTGTGGCATTGACGAGTGGTGGAGGAATGAACAGTTCCTGGGTCAATGGAGGTATC 2855
Db |||||
QY 2971 AGCGCGTGGGCATCGACGAGTGGTGGAGGAACGAGCAGTCTTGGGTCACTGGAGGCATC 3030
QY 2856 TCTGCACATCTGTTTGGCGTCTTTCAGGGTCTTTCAGAGTCTTTCAGAGTCTGCGGTATCGACACC 2915
Db |||||
QY 3031 TCCGCCACCTCTTTCGCCGTCTTTCAGGGCTCTTTCAGAGTCTTTCAGAGTCTGCGGTATCGACACC 3090
QY 2916 AACTTCACTGTCACCTCAAGGCTAATACGAGAGAGCGGCTTTCAGTCTGAGCTCTACATG 2975
Db |||||
QY 3091 AACTTCACTGTCACCTCAAGGCTTTCAGAGAGAGCGGCTTTCAGGAGTCTGAGTCTGATG 3150
QY 2976 TTCAAGTGGAGCGGCTTCTCATCTCCCTCCGACGACCATTTTGGATCAATTAACATGTTGGT 3035
Db |||||
QY 3151 TTCAAGTGGAGCGACTTCTGATCCCGCCACCATCTCTGATCATCACTGTCGCGC 3210
QY 3036 GTCGTTGCTGGACCTCTCAGCCATCAACAGTGGTTACCAATCATATGCGGCGCGCTCTTT 3095
Db |||||
QY 3211 GTTGTGCGCGCATCTCTCAGCCATCAACAGCGGGTACCAGTCTGTCGGGTCTCTCTC 3270
QY 3096 GGGAGGCTCTCTTCTGCTCTGCGGTGATCTTCACTTATACCATCTCTCAAGGCTCTT 3155
Db |||||
QY 3271 GGCAGGCTCTTCTTCTGCGCTTCTGGGTGATCTGTTCACTGTATCCCGGCTCTCAAGGCTCTC 3330
```

```
QY 3156 ATGGGAGGCAAAACCCGACACCGACGATTTGTCATCTGCTGGGCTGTCCTCTCGCTTCT 3215
Db |||||
QY 3331 ATGGGTCGGCAGAACCGACCCCGACCATCGTGTGTTCTCTGGGCGATCTCTGCTGGCTCG 3390
QY 3216 ATCTTCTCTCTGCTGCTGGGTTGCTGTGATCCATTCACCTACCGTCTCGCTGGCCCAAT 3275
Db |||||
QY 3391 ATCTTCTCTCTGCTGCTGGGTTGCTGTGATCCATCGATCCGTTTACCACCGCTCTCTG 3450
QY 3276 ATCCAACTGTCGCTCACTCACTAGGAAAGTGGGA-GTTTCTAGAGACAGAAA 3329
Db |||||
QY 3451 ACTCGAACGTCGTCATCACTCTAGGAGGTGGAAGGTTTGTAGAACAGAGA 3505

RESULT 6
AF200528 3745 bp mRNA linear PLN 31-AUG-2000
LOCUS Zea mays cellulose synthase-4 (Cesa-4) mRNA, complete cds.
DEFINITION AF200528
ACCESSION AF200528
VERSION AF200528.1 GI:9622879
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3745)
AUTHORS Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S.,
Xocoostle-Cazares, B., and Delmer, D.P.
TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene
family
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3745)
AUTHORS Dhugga, K.S. and Helentjaris, T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES
Location/Qualifiers
1..3745
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/note="Identified and isolated from the maize genome
database cDNA collection of Pioneer Hi-Bred International,
Inc"
1..3745
/gene="Cesa-4"
320..3553
/gene="Cesa-4"
/note="CESA-4"
/codon_start=1
/product="cellulose synthase-4"
/protein_id="AAF89964.1"
/db_xref="GI:9622880"
```

```
/translation="MEGDAGVKSRRGGQVQICGDSVGTAAEGDVFAACDVGFP
VCRPYERKDGTCACQCKTKYKHKGSPAIRGEGDDTDADSDNFYASGNEDOK
QKADMRSMRVNMGSGDVGPRKYDGEIGLTKYDSGEIPRGYIPSVTNSQISIEIP
GASPDHMMSPGTNIGKRAPPPYVNHSPNPSRPSGSIENVKERVDSQWKKDKGT
IPWTCGSIAPSEGRGVDIDASTDYNMEDALLNDETRQPLSRKVPDPSSRINRPMV
IVLRLVLSIFLHYITNVRNAYPLMLLSVCEIWFALSWILDQFFKFPINRETYL
DLRLRYDRGEPSSQLAADVFVSTVDPMEPLVNTANTVLSILAVDPYDVKVSVYS
DGAAMLTFDALAEISFARKWPFVVKYNIENRAPWYFSQKIDYLDKWKHPSFVKD
RRMKREYEFKRVNGLVAKAOKVPEGMIMODGTPWGNNTDHPGMQLQVFLHSG
GLDTEGNELRVLVVSREKRPQHKKKAGAMNALVRVSAVLNQCMLNDCDHYIN
NSKALREAMCLMDPNLGRSVCIYQKQFQFDGLDRNDRNRYANRNTVFFDINLRGLDGIQ
GPYVGTGCFNRTALYGPPIPKKGGFGLSSLCGGRKXASKSKSKSKSKSKSKQKQHD
SSVPFNLEDEEGVEGAGDEKSLMSQLEKFGQSAAFVATLMRYGVSQVPA
PLTSLKLEAIIHVSQGVDETEWGTETIGTYSVTEDILTGFKNHARGWRSIYCMPR
PAFGSAPINLRLNQLRWALGSVEILFSRHCPMLWYGVGGRKLTLEFAYINTLY
PLTSLPILLYCILPAICLLTKETLPEISNFSASIFLSIFATGILBMRWGSVGI
DEWRNEQFVIGGISNHLFAVCGLLKVLGADITNFTVTSKASDEGDFAEFLYFMW
TTLILPTTILILINLVGVAGISYAINSGYSQSWGPLFGKLFAPFVIVHLVYPLFKGLM
```

GRQNRPTIVVWAILLASIPSLWVRIDPFTTRVTRVGPDTQTCGINC"

ORIGIN

Query Match 70.9%; Score 2572.2; DB 8; Length 3745;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

QY	4	CGAGGAACCCCGCTCAGACTCTGTCTGTCGGTGGGTTGGATCGCTCGCGCGGCATGG	63
DB	264	CCAGGCCCCAGGCTCCAGGCCAGCTCCCTCGACGTTTTCGGCGAGCTCGCTGGCCATGG	323
QY	64	ACGGCCACCGGACGCCCTGAACTCCGGGAGGCACCGGGCGCGGGAGCGTGTGCCAGATCT	123
DB	324	AGGGCCACCGGACGGGCTGAAGTCCGGGAGGCGCGGTGGCGGACAGGTGTGCCAGATCT	383
QY	124	GCCCCGACGCGCTGGGACCCACAGCTTGGACGGCGACGCTTTCACCGCTCGGACGTCTGCC	183
DB	384	GCGGCGACGCGCTGGGACCCACCGCGGAGGGGACGCTTTCGCGCGCTCGGACGTCTGGC	443
QY	184	GCTTCCCGTCTCGCGCCCTGCTAGCAGCACGAGCGCAGGAGGACCCAGCGCTTGC	243
DB	444	GGTTCCCGTGTCCCGCCCTGCTACGAGTACGAGCGCAAGGACGGCACCGAGCGTGC	503
QY	244	TCCAGTGCAGAACCAAGTACAAGCGCCACAGAGGAGCGGCCAGCGATCCGCGGGGAGGAAG	303
DB	504	CCAGTGCAGAACCAAGTACAAGCGCCACAGGAGGAGCGGCCAGTCCGTTGGGGAGGAAG	563
QY	304	GCAGCACACTGATGCCGATGATGTAGTGACTTCAACTACCTGCACTTGGCACTGAGG	363
DB	564	GAGACGACACTGATGCCGA-----TAGCGACTTCAATTACCTTGGCATCTGGCAATGAGG	617
QY	364	ACCAGAGCAGAGATTTCTGACAGGATGCGAGCTGGCGCATCAACACCGGGGCGAGTG	423
DB	618	ACCAGAGCAGAGATTTCCGACAGAAATGCGAGCTGGCGCATCAACGTTGGGGGCGAGG	677
QY	424	GCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAGTATGACAGTG	483
DB	678	GGGATGTTGGTGGCCCCCAAGTATGACAGTGGCGAGATCGGCGCTTACCAAGTATGACAGTG	737
QY	484	GAGAGATCCCTAGGGGATAGCTCCCTTCAGTACACCAACAGCGAGATGTAGGAGAAATCC	543
DB	738	GCGAGATTCCTCGGGGATACATCCATCAGTCACTAACAGCGAGATCTCAGGAGAAATCC	797
QY	544	CTGAGACTTCCCTGATCATCATGATGTCCCTTACGGGAAACATCAGCAGAGCTGCTC	603
DB	798	CTGGTGTCTCCCTGACCAATCATATGATGTCCCACTTGGGAAATTTGGCAAGCGTCTC	857
QY	604	CGTTCCCTATGTGAATCATTCACCAATCCGTCAAGGGAGTTCTCCGGCAGTATTGGGA	663
DB	858	CATTTCCTATGTGAACCATTCGCCAATCCGTCAAGGAGTTCTCTGGTAGCATTTGGGA	917
QY	664	ATGTTGCTGGAAGAGAGAGTGTGATGGCTGGAAATTAAGCAGGACAGAGGTGCGATTG	723
DB	918	ATGTTGCTGGAAGAGAGGTTGATGGCTGGAATTAAGCAGGACAGAGGTGCGATTG	977
QY	724	CCATGACTAATGGACAGCATTTCTCCCTCTGAGGTGCGGCGAGTACTCAGCATCGATG	783
DB	978	CCATGACGAATGGACAGCATTTCTCCCTCTGAGGGTGGGGTGTGGTGAATTTGATG	1037
QY	784	CATCTACTGAATACAACTGGAAGACGCTTTACTGAATGATGAATCTGCCAGCGCTCTAT	843
DB	1038	CATCAACTGATTAACAATGGAAGATGCTTTATTGAACGACGAACTCGACAGCTCTAT	1097
QY	844	CTAGAAAGTCCCATTCCTTCTCCAAATAAATCCCTCAGAAATGCTCAATTTGTTCTGC	903
DB	1098	CTAGAAAGTCCCATTCCTTCTCCAGGATAAATCCATACAGGATGCTCAATTTGTTCTGC	1157
QY	904	GTTGGTGTCTAAGCATCTTCGCACTACCGTCTCACAATCTCGTGTGCTAATGAT	963
DB	1158	GATTGATTTCTAAGCATCTTTGCACTACCGTATCAAAATCTGTGTGCGAATGAT	1217
QY	964	ACCCACTGTGGCTTTATCTGTATATGTGAGATTTGGTTGCTTTATCTCTGGATCTGG	1023

DB	1218	ACCATTATGGCTTCTATCTCTTATATGTAGATCTGGTTTGTCTCTTTCGTGTGATATTGG	1277
QY	1024	ATCAGTTCCTCGAAGTGGTTTCCAAATCAACGGGAGACCTACCTTGTATAGACTGGCTTTAA	1083
DB	1278	ATCAGTTCCTTAAGTGGTTTCCAAATCAACGGGAGACCTACCTTGTATAGACTGGCTTTAA	1337
QY	1084	GGTATGACCGGAGAAGGTGAACCGTCTCAGTTGGTCTCTGTGACATATTGTTCAGTACAG	1143
DB	1338	GGTATGACCGGAGAAGGTGAGCCATCTCAGTTGGTCTCTGTGACATATTGTTCAGTACAG	1397
QY	1144	TCGACCCCTTGAAGGAGCCACTATCTGCTCACTGCCAAGCTGTGCTATCCATTTCTGCTG	1203
DB	1398	TCGACCAATGAAGGAGCCCTCTCTTGTCACTGCCAATACCGTGTCTATCAITTTCTGCTG	1457
QY	1204	TTGATTTATCCGTTGGACAAGGCTCTCTTGTATGTATCTATGATGACGAGCTTCAATCTGA	1263
DB	1458	TGATTTATCCCTGTGGATAAGGCTCTTGTCTATGTATCTGATGATGAGCTGCGATCTGA	1517
QY	1264	CTTTTGAACGATTTGGCTGAGACTTCAAGTGTCTAGGAAAATGGGTACCATTTTGTGAAGA	1323
DB	1518	CATTTCGATGCACTAGCTGAGACTTCAGAGTTTGTCTAGAAAATGGGTACCATTTTGTGAAGA	1577
QY	1324	AGTATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTCCAGAAAAATTTGATTAAGCTGA	1383
DB	1578	AGTATGACATTTGAACCTAGAGCTTCTGAAATGGTACTTCTCCAGAAAAATTTGATTAAGCTGA	1637
QY	1384	AAGACAAAGTCCAGGCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATTAAG	1443
DB	1638	AGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATTAAG	1697
QY	1444	AATTTAAATTCAGATAAATGCCCTAGTTTCTAAGGCAITTTGAAAGTCCCGAGGAGGAT	1503
DB	1698	AATTTAAATTTAGGTTAAATGGCTTGTCTAAGGCACAGAAAATTTCTTGAGGAGGAT	1757
QY	1504	GGATCATGCAAGATGGCACACCATGCGCCAGGAACATACACGAGGATCACTCTGGAATGA	1563
DB	1758	GGATCATGCAAGATGGCACACCATGCGCCAGGAACATACCMGGACCATCTCTGGAATGA	1817
QY	1564	TTCAGGTTTCTTGGTTCAGGTGGGCTTGTATCTAGGGGTAAATGAGCTCCCGGTT	1623
DB	1818	TTCAGGTTTCTTGGTTCAGGTGGGCTTGTATCTAGGGGCAATGAGCTACCGGTT	1877
QY	1624	TAGTTTATGTCTCTGGAAGAGGCTCTGGTTCCAGACCAACAGAGGCTGGTGGCA	1683
DB	1878	TGGTCTATGTTTCTCGTGAAGAGGCTCTGGAATTTCCAGCATCAACAGAAAGCTGGTGGCA	1937
QY	1684	TGAATGCTTCTTGGTTCAGGTGGGCTTCTTACTAATGACATATCATGTTGAATCTTG	1743
DB	1938	TGAATGCTTCTTGGTTCAGGTGGGCTTCTTACTAATGACATATCATGTTGAATCTTG	1997
QY	1744	ATTGTGATCACTACATCAACAGCAAGGCTGTCCGAGAAGCTATGTGTTCTTAATGG	1803
DB	1998	ATTGTGATCACTACATTTAAACAGTAAGGCTCTCAGGAGGCTATGTGTTCTTATGG	2057
QY	1804	ATCCAAACCTAGGTCGCGAAGTCTGTTATGTGAGTTCCCAAGGTTTGTATGGGATTG	1863
DB	2058	ACCCTAACCTAGGAAGGAGTGTCTGTACGTCCAGTTTCCCGAGAGTTCCGATGGATG	2117
QY	1864	ATAGGAATGATCGATATGCAACAGGAACACTGTTCTTTTGTATTAATTAACCTTGAGGGCC	1923
DB	2118	ACAGGAATGATCGATATGCAACAGGAACACTGTTCTTTTGTATTAATTAACCTTGAGGGCC	2177
QY	1924	TTGACGGCATTCAGAGGACAGTTTATGTGGAACTGGTTGTGTTTTTCAACAGAAACAGCTA	1983
DB	2178	TTGATGGCATTCAGAGGACAGTTTATGTGCGAACTGGGCTGTTTTTCAACAGAAACAGCTC	2237
QY	1984	TCTATGTTATGACCCGCCAATTAAGGGAAGAGCCAGGTTTCTTGGCATCACTATG	2043
DB	2238	TATATGTTATGACCCGCCAATTAAGGGAAGAGGTTGTTCTTGTTCATCACTATG	2297
QY	2044	GGGCGAAGAAAGCAAGCAAGCTCAAGAAAAAGGAGCTCAGATAAGAAAAAGTCCGAA	2103
DB	2298	CGGTAGGAAGCAAGCAAGCAAGCTCAAGAAAAAGGAGCTCAGATAAGAAAAAGTCCGAA	2354

Qy	2104	AGCATGTGGACAGTCTCTGTCTCCAGTATTCAATCTCGAAGACATAGAGGAGGGTCTTGAAG	2163
Db	2355	AGCATGTGGACAGTCTCTGTCTCCAGTATTCAACCTTTGAAGATATAGAGGAGGAGTCTGAAG	2414
Qy	2164	GTGCTGGGTTTGATGATGAGAAATCAGTTCTCTCATGTCTCAATCAGCTTTAGAGAAGAGAT	2223
Db	2415	GGCTGGATTGACGACGAGAAATCACTTCTTATGTCTCAATGAGGCTTGGAGAAGAGAT	2474
Qy	2224	TTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTCTTCCCTCAGT	2283
Db	2475	TTGGCCAGTCGCGCAGCGTTTGTGCCCTCCACTCTGATGGAGTATGGTGGTCTTCCCTCAGT	2534
Qy	2284	CCCTCACTCCAGATCTCTTTTGAAGAGCATATCCATGTCTAATGATCTGTGCGCTATGCG	2343
Db	2535	CCGCACTCCGGAGTCTCTCTGAAAGAGCATATCCATGTTTATAAGCTGTGGCTATGAGG	2594
Qy	2344	ACAAGTCTGAATGGGGAACCTGAGATTGGTTGGATCTATGGATCTGTGCACAGAAGATATTC	2403
Db	2595	ACAAGACTGAATGGGGAACCTGAGATCGGTGGATCTACGGTTCGTGTACAGAAGACATTC	2654
Qy	2404	TTACTGAAATCAAGATGCGACGAAGAGCTTGGCGTTTCACTATATGATGCCAAGCGCC	2463
Db	2655	TCACCGGATTCAAGATGACGCGCGAGGCTGGCGGTGCACTACTGTCATGCCAAGCGGC	2714
Qy	2464	CAGCTTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGCTGTGAACCAAGTCTGCGGT	2523
Db	2715	CAGCTTTCAAGGGTCTGCCCCCATCAATCTTTTCGACCGTCTGAAACAGGTCTCCGGT	2774
Qy	2524	GGGCTCTCGGTTCTGTGTGAAATCTTTTCAGCCGGCATATGCCCTTATGTTATGGCTACG	2583
Db	2775	GGGCTCTGGTCCGTGGAGATCCTCTTCAGCGGCACTGCCCTCTGGTACGGCTACG	2834
Qy	2584	GAGGGCCCTCAAGTTCCTCTGAGAGATTCGGTTTACATCAACACACCAATTTACCCACTAA	2643
Db	2835	GAGGGCGGCTCAAGTTCCTGAGAGATTCGCGTACATCAACACCAACCATCTACCGCTCA	2894
Qy	2644	CCCTCTCCCGCTCTCTAGTCTATTGTATATTGCTGCTATCTGCTCTGCTCACTGGAAAGT	2703
Db	2895	CGTCCATCCCGCTTCTCATCTACTGCAATCTGCGCCGCATCTGCTCTCACCGGAAAGT	2954
Qy	2704	TCATCATGSCAGAGATTAGCAACTTGGCCAGTATCTGGTTCAATGCGCTCTTCTCTTCAA	2763
Db	2955	TCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTCACTCCCTCTTTCATCTCGA	3014
Qy	2764	TTTTGCGCACTGGTATCTTTGAGATGAGGTGGAGTGGTGTGGCAATGACGATGGTGGGA	2823
Db	3015	TCCTTGGCCACGGGCATCTCTGGAGATGAGGTGGAGCGGGGTGGGCATCGACGAGTGTGGGA	3074
Qy	2824	GGAAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTCAGG	2883
Db	3075	GGAAACGACAGTTCTGGGTGATCTGGGGGCACTCTCCGGGCACTCTTCGCGGTGTTTCAGG	3134
Qy	2884	GTCTTCTGAAGGTGCTTGGCCGATCGACACCAACTTCACTGTCACTCAAGAGGCTAATG	2943
Db	3135	GCCTGTCTCAAGTGTGCGCCGGCATCGACACCAACTTCAACCGTCACTCAAGGCTCTCG	3194
Qy	2944	ACGAGAGCGGACTTTGCTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCCTC	3003
Db	3195	ACGAGGACGGCACTTCGCGGAGCTGTACATGTTTCAAGTGGAGAGCGTCTCTGATCCCGC	3254
Qy	3004	CGACGACATTTTGATCATTTAAATGTTGGTGTGGTGTGGCACTCTCTACGCGCATCA	3063
Db	3255	CCACCACCATCTGATCATCAACTGGTGGGCTGTGCGCGGCACTCTCTACGCGCATCA	3314
Qy	3064	ACAGTGTTTACCAATCATATGGGGGCGCTCTTTGGGAAGCTCTTCTTTGGCTTCTGGGTGA	3123
Db	3315	ACAGCGGATACCAAGTCTGTGGGGCCGCTCTTCGCAAGCTCTTCTTCGCTTCTGGGTCA	3374
Qy	3124	TTGTTTCACTTATACCAATCTCTCAAGGGTCTTATGGGACGGCAAAACCGCACACGACGA	3183
Db	3375	TCGTGCACTGTATCCGCTTCTCAAGGGCTCTATGGGACGAGAAACCGCACCCCAACCA	3434

Qy	3184	TTGTCATCGTCTGGCGTGTCTCTCTCGTCTCTATCTTCTCTCTGTGCTGTGGTTCGTGTG	3244
Db	3435	TGCTGCTGCTGTGGCCCATCTCTGTGCCTGCCATCTTCTCTGCTGTGGTTCGCATCG	3494
Qy	3244	ATCCATTCACTACCCGTCCTCGCTGGCCAAATATCAAACCTGTGGCATCAACTGCTAGG	3303
Db	3495	ACCCTTTCAACCCCGCTCACTGGCCCGATACCAGACGTGTGGCATCAACTGCTAGG	3554
Qy	3304	AAAAGTGGGAGTTT	3316
Db	3555	GAAAGTGAAGTTT	3567
RESULT 7			
AK069196			
LOCUS			
DEFINITION	AK069196 4282 bp mRNA linear PLN 24-JUL-2003		
ACCSSION	Oryza sativa (japonica cultivar-group) cDNA clone:J023003G18, full insert sequence.		
VERSION	AK069196.1 GI:32979220		
KEYWORDS	FLI cDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
AUTHORS	1		
	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team.; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaka,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sakaki,D., Sato,K., Shibata.K., Shinagawa,A., Shiraki,T., Yoshino.M. and Hayashizaki,Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science 301 (5631), 376-379 (2003)		
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2 (bases 1 to 4282)		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Oeato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sakaki,D., Sato,K., Satoh,K., Shibata.K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano.S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi.F., Takaku-Akaira S., Tanaka.T. Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi.W., Yamada,H., Yamamoto,M., Yasunishi.A., Yazaki,J., Yokomizo,S. and Yoshimura,A.		
Direct Submission			
Submitted (05-DEC-2001)	Shoehi Kikuchi. National Institute of		
JOURNAL			

1803 ATGAGCTTCCTCGTTAGTCTATGTTGCTCGTGAGAACGTCCTGGATTCACACCAATA 1862
1669 AGAAGGCTGGTGCATGAATGACCCCTGTTGGTGTCTCAGCTGTCTCTTACTAATGACCAAT 1728
1863 AAAAGGCTGGTGTATGAATGACCTGTTGCTGTATCAGCTGTCTCTTACTAATGACCAAT 1922
1729 ACATGTTCAATCTGATTTGATGATCACTACATCAACAAAGCAAGCAAGCTGTCCGAGAGCTA 1788
1923 ACTTGTGAAATCTTGACCTGTGATCACTACATCAACAAAGCAAGCAAGCTGTCCGAGAGGCTA 1982
1789 TGTGCTTCTTAATGATCCAAACCTAGGTCGCGCAAGTCTGTTATGTCAGAGTTCCACAAA 1848
1983 TGTGCTTCTTAATGATCCAAACCTAGGTCGCGCAAGTCTGTTATGTCAGAGTTCCACAAA 2042
1849 GGTGTTGATGGGATGATAGGAATGATCGATATGCAAAAGCAAGCAAGCTGTCTCTTTTGTGATA 1908
2043 GGTGTTGACGGTATCGATAGAAATGATCGATATGCAAAAGCAAGCAAGCTGTCTCTTTTGTGATA 2102
1909 TTAACCTTGAGGGGCTTTCAGCGCATTCAGGACCAAGTTTATGTCGGAACCTGTTGTTGTT 1968
2103 TTAATTTGAGAGGCTTCGATGCTTCCAGGACCAAGTTTATGTCGGAACCTGTTGTTGTT 2162
1969 TCAACAGAACAGCTATCTATGTTATGAGCCGCCCAATTAAGGCGAAGAGCCAGGTTTCT 2028
2163 TCAACAGAACAGCTATCTATGTTATGAGCCGCCCAATTAAGGCGAAGAGCCAGGTTTCT 2222
2029 TGGCATCATATGTTGGGGCGAAGAAAGAGGCAAGCAAGTCAAAAGAAAGAGGAGCTCAGATA 2089
2223 TCTCTTCGCTTTGTTGGGGCGAAGAAAGAGCAAAAGAGTCAAGGAGAGAGCAAGGAAA 2282
2089 AGAAAAAGTCGAAACAGCATGTGGACAGTCTGTTCCAGTATTCATCTCGAAGACATAG 2148
2283 AGAAAAAGTCGAAACAGCATGTGGACAGTCTGTTCCAGTATTCATCTCGAAGACATAG 2342
2149 AGGAGGCTGTTGAAGGCTGCTGGGTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGA 2208
2343 AGGAAGGATTTGAAGTTCTGGATTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGA 2402
2209 GCTTAGAAGAGATTTGGCCAGTCAGGAGCAATTTGTCCTCCACTCTGATGGAATATG 2268
2403 GCTTAGAAGAAAGATTTGCTCAATCTAGTGTGTTTGTAGCCTCCACTCTGATGGAATATG 2462
2269 GTGGTGTTCCTCAGTCCTCCACTCAGATCTCTTTTGAAGAAGCTATCCATGTCATAA 2328
2463 GTGGTGTTCCTCAATCCGCAACTCAGAAATCTTTCTGAAAGAGGCCATACATGTTATCA 2522
2329 GTTGTGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGGTGGATCTATGATCTG 2388
2523 GCTGTGGCTATGAAGACAAAGTGAATGGGAACTGAGATTTGGATGGATCTATGTTCTG 2582
2389 TCACAGAGATATTTCTTACTGATTCAGATGCGACGCAAGAGGCTGGGCTCAGTCTATT 2448
2583 TTACAGAGATATTTCTACTGATTCAGATGCGATGCGATGCGGCTGGGCTCAATCTACT 2642
2449 GCATGCCCAAGCGCCAGCTTTCAGGATCTGCCCCCATCAATCTTTCAGATGCTCTGA 2508
2643 GCATGCCCTAACGACCAATCAAGGGGCTGCCCCCAATTAATCTTTCGATGCTCTGA 2702
2509 ACCAAGTGTCTCGGTGGGCTCTCGGTCTGTTGTTGAAATCTTTTCAGCGGCAATGGCCCT 2568
2703 ACCAGTGTCTCGGTGGGCTCTCGGTCTGTTGTTGAAATCTTTTCAGCGGCAATGGCCCT 2762
2569 TATGTTATGGCTACGAGAGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACCA 2628
2763 TATGTTATGGCTACGAGAGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACCA 2822
2629 CCATTACCCACTACCTCTCCGCTTCTAGTCTATGTTATGTTATGTTGCTGCTATCTGTC 2688
2823 CTATTATCCACTACATCGATTCACCTCTGTTGATGATGATGATGATGATGATGATGATGAT 2882
2689 TGCTCAGTGGAAAGTTTCAATCATGCCAGAGATTAGCAATTTGGCCAGTATCTGTTTCAATTG 2748

Db 2883 TTCTCACTGGGAAGTTTCATCATACAGAGATTAGCAACTTTTCGAGATTTTGGTTTATCT 2942
QY 2749 CGCTCTTCTTCCATTTTCGCAATTTTCGCACTGTGTATCTTTGAGATGAGGTGGAGTGTGTGGCA 2808
Db 2943 CTCTCTTCTGTCAATCTTTGCTACTGTGTATCTTTGAGATGAGGTGGAGTGTGTGGCA 3002
QY 2809 TTGAGAGTGTGGAGGAATGACAGTTCTGGGTCAATTTGGAGGTATCTCTCCACATCTGT 2868
Db 3003 TCGATGAATGTGGAGGAATGAGCAGTTTGGGTTATTTGGAGGTATTTCCGCTCATCTAT 3062
QY 2869 TTGCGCTCTTTTCAGGCTCTTCTGAGGTGCTTTGCGGTATCGACACCAACTTCAGTGTCA 2928
Db 3063 TTGCGGCTCTTCAGGCTCTGCTGAGGTGCTTTGCTGGAATTTGATACCAAGCTTCATGTCA 3122
QY 2929 CCTCAAGGCTTAATGACGAAAGAGGCGACTTTGCTGTGAGCTCTACATGTTCAAGTGGAGCA 2988
Db 3123 CCTCAAGGCGCTCTGATGAAGAAGGCGATTTGCGTGTGAGCTCTACATGTTCAAGTGGAGCA 3182
QY 2989 CGCTTCTCATCCCTCCGACGACCAATTTTGTATCATTAACATGTTGCTGTGCTGGCA 3048
Db 3183 CCCTTCTGATCCCAACCGACCACTATTTTGTATCATCAACTGCTGTGCTGTGTTGCTGGTA 3242
QY 3049 CCTCTACGCCATCAACAGTGGTTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTTCT 3108
Db 3243 TCTCTACGCTATCAACAGTGGTACCACTCATGGGCTCGCTCTTTGGGAAGCTCTTCT 3302
QY 3109 TTGCTTCTGGGTGATGTTTCACTTATACCATTCCTCAAGGCTCTTATGGGAGCGGAAA 3168
Db 3303 TTGCTTCTGGGTGATGTTTCACTTATACCATTCCTCAAGGCTCTTATGGGAGCGGAAA 3362
QY 3169 ACCGACACCCAGCATGCTCATGCTGGGCTGCTCTCTCGTCTCTATCTCTCTCTGTCG 3228
Db 3363 ACCGACACCCAGCATGCTCATGCTGGGCTGCTCTCTCGTCTCTATCTCTCTCTGTCG 3422
QY 3229 TGTGGGTTCTGTTGATCCATTCATCCGCTCTCGCTGGGCCCAAAATATCCAAACCTGTG 3288
Db 3423 TGTGGGTTCTGATCGATCCATTCACCCCGCTGTCACAGGCGCTCATACCCAAAATGTG 3482
QY 3289 GCATCAACTGCTAGAAAGTGGGAGTTTGTAGAGACAGAA 3328
Db 3483 GTATCAACTGCTAGAGTAGAGCAAGATAGTAGGCAAGAA 3522

RESULT 8
AX652952
LOCUS AX652952 3222 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 2822 from Patent WO03000898.
ACCESSION AX652952
VERSION AX652952.1 GI:29155766
KEYWORDS
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 2822 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source location/Qualifiers
1..3222
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Query Match 64.9%; Score 2352.4; DB 6; Length 3222;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 2675; Conservative 0; Mismatches 516; Indels 3; Gaps 1;
QY 109 ACGTGTGCCAGATCTGCCCGACGCGCTGGGCACACAGTGTGGAGCGGACGCTCTTCACCG 168

Db 377 TTCCCGTGTGCGCCATGCTACGAGTACGAGCGCAAGCGCGTCCCG 436
Qy 246 CAGTGCAGACCAAGTACAGCGCCACAGAGGAGCCAGCGATCCGCGGGAGGAGGC 305
Db 437 CAGTGCAGACTAAGTACAGCGCCACAAAGGAGCCACAGTACACGCTGAGGAAAT 496
Qy 306 GAGGACACTGATCCGATGATGTAGTACTCACTACCTGCACTGCACTGAGGAC 365
Db 497 GAGGATGTGATGCTGACGATGTAGTACTCACTACCACTACCAAGCATCTGCAACACCGGAT 556
Qy 366 CAGAAGCAGAAGATGTGACAGGATGCGCAGCTGCGCATGAACACCGGGGCGAGTGGC 425
Db 557 CAGAAGCAAGANTGCTGAGAGATGCTCACTTGGGAGCAACTCACGTGCGCATG-- 614
Qy 426 AATGTTGGCCACCCCAAGTATGACAGTGGGAGATCGGCTCTCCAAAGTATGACATGGA 485
Db 615 -ATATTTGGCGTGGCTAAGTATGACAGCGGTGAAATTTGGGCATGGGAAGTATGACAGTGGT 673
Qy 486 GAGATCCCTAGGGATACGTCCTTCAGTCACCAACAGCCAGATGTGAGGAGAAATCCCT 545
Db 674 GAGATCCCTGCTGGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCT 733
Qy 546 GAGGCTTCGCTCATCATCATGATGCTCCCTACCGGGAACATCAGCAGACGTGCTCCG 605
Db 734 GAGGCTTCCTCATCATATGATGCTCCTGTTGGGAACATTTGGCGAGCGGTGACATCAA 793
Qy 606 TTTCCTATGTGAATCATTCACCAATCCGTCAAGGAGTCTCCGGCAGTATTTGGGAAT 665
Db 794 TTTCCTATGTAAATCATTTCCAAACCCATCAGGAGGTCTCCCGTAGCCTTGGCAAT 853
Qy 666 GTTGCTTGGAAAGAGAGATTTGATGCTGGAAATGAAGCAGCAGCAGGCTGCGATTCCC 725
Db 854 GTTGCTTGGAAAGAGAGGTGATGATGGAATGA--AGGATAAGGTGCAATTCCT 910
Qy 726 ATGACTAATGGGCAAGCATTTGCTCCCTCTGAAGGTGGGAGCTACTGACATCGATGCA 785
Db 911 ATGACCAATGGAACAAGCATTTGCTCCATCAGAAGGCGTGAGTGTGCTGATATGATGCT 970
Qy 786 TCTACTGAATACACATGGAAGAGCTTTACTGAATGATGAATCGCCAGCTCTATCT 845
Db 971 TCTACTGAATATACATGGAAGATGCTTACTGAATGATGAATCGGCAACCTCTATCT 1030
Qy 846 AGAAAGTCCCATTTCTCCCAATAAATCCCTACAGATGTCATTTCTGCGG 905
Db 1031 AGAAAGTCCCATTTCTCCATCCAGNATAATCCGTCAGAAATGTCATTTGCTACGT 1090
Qy 906 TTGGTTGTTCTAAGCATTTCTCCACATCCGCTCTCAAAATCCTGTGCGTAATGCATAC 965
Db 1091 TTGGCTGTTCTATGCATATTTCTTGGCTACCGTATCACACATCCTGTGAACAATGCATAT 1150
Qy 966 CCACTGTGGCTTTATCTGTATATGTGAGATTTGGTTGCTTTATCTCGGATCTGGAT 1025
Db 1151 CCACTGTGGCTTTATCCGTCATATGTGAGATCTGGTTGCTTTGCTTGGATTTTGGAT 1210
Qy 1026 CAGTCCCGAAGTGTTCACATCAACCGGAGACCTACCTTCATAGACTGGCTTTAAGG 1085
Db 1211 CAGTCCCAAGTGTGTCCTCAATCAACCGTGAACATACCTTCATAGACTGGCTTTAAGG 1270
Qy 1086 TATGACCGAAGGTGAACCGCTCTCAGTTGGCTGCTTTGACATATTTGTCAGTACAGTC 1145
Db 1271 TATGACCGAAGGTGAACCATCTCAATTAGCTCTCTGATATTTTGTGCACTACTGTG 1330
Qy 1146 GACCCCTTGAAGAGCCACTATCTGCTACCTGCAACACTGTGCTATCCATTTCTGCTGT 1205
Db 1331 GATCCAAATGAAGAGCGCTCTCTGTGCTACCTGCAAAATCTGTGCTTCCATCTTGTGTC 1390
Qy 1206 GATTATCCCGTGAAGGTGCTCTTGTATGTATCTGATCAGCGAGCTTCAATGCTGACT 1265
Db 1391 GATTATCCCGTGAAGGTGCTCTTGTATGTATCTGCTATGTTTGGATGATGAGCTGCTATGCTGACT 1450
Qy 1266 TTTGACGCATTTGGCTGAGCTTCAAGTTTGTAGGAAATGGGTACATTTGTTGAAGAAG 1325

Db 1451 TTTGATGCTCTCTCTGAAACTTTCAGAGTTTGTAGAAAAATGGGTTCCGTTCTGTAAAG 1510
Qy 1326 TATGACATTTGAACCCAGAGCTCCGAGTTTATCTTTTGCAGAAAAATGATTAACCTGAAA 1385
Db 1511 TACAACATAGAGCTTANGCCCGGAAATGGTACTTTGCTCAGAAAAATGATTAACCTGAAA 1570
Qy 1386 GACAAAGTCCAGCTTCAITTTGTTAAAGACCGCGGCGCATGAAGAGAAATATGAAGAA 1445
Db 1571 GACAAAGTTCAAACCTCATTTTGTGAAGAAACGCGGCGCATGAAGAGAAATATGAAGAA 1630
Qy 1446 TTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTTGAAAGTCCCGGAGGAGATGG 1505
Db 1631 TTTCAAAGTTTCGTATCAATGCTCTTGTAGCCAAAGGCAAAAAAGTTTCCGAGGAGGATGG 1690
Qy 1506 ATCATCAGATGGCACACATGCGCCAGGAAACAATACAGGATCATCTCGGAATGATT 1565
Db 1691 ATCATCAGATGGTACACCTTGGCTGGGAAACAATACAGGATCATCTCGGAATGATT 1750
Qy 1566 CAGTTTTCCTTGGTCACAGTGGTGGCTTGTATCTACGAGGTAAATGAGCTCCCGCTTTTA 1625
Db 1751 CAGTTTTCCTGGTTCACAGTGGAGGCTTGAAGTTGAAGGCAATGAATTCCTCTGTTG 1810
Qy 1626 GTTATGTGCTGCTGTAAGAGCTCTCGGTTTCAGACCAACAAGAGGCTGTGCGCATG 1885
Db 1811 GTTATGTGCTGCTGTAAGAGAGCTCTCGGATTCGAATTCGAATTCGAATTCGATG 1870
Qy 1686 AATGCCCTTGTGCTGCTCTCAGCTGCTTACTTAATGGAACAATACATGTTGAATCTTGAT 1745
Db 1871 AATGCCCTTGTGCTGCTGCTGCTTACTTAATGGAACAATACATGTTGAATCTTGAT 1930
Qy 1746 TGTGATCACTACATCAACAACAGCAGGCTGTCCGAGAGCTATGTGCTTCTTAATGGAT 1805
Db 1931 TGTGACCATCATCATCAATTAAGCAAGGCTCTTCGAGAGCTATGTGCTTCTTATGGAC 1990
Qy 1806 CAAAACCTAGTCCGCAAGTCTGTATGTGAGTCTCCCAAAAGGTTGATGGATGAT 1865
Db 1991 CAAAACCTAGGAGGATGTGTTATGTGCAATTTCTCAGAGGTTGATGGATGAT 2050
Qy 1866 AGGAATGATCGATATCAACAACAGGACACCTGCTTTTGTATTAATTAACCTGAGGCGCTT 1925
Db 2051 AGGAATGACCGATATGCAACAACAGGACACCTGCTTTTGTATTAATTAACCTGAGGCTT 2110
Qy 1926 GACGGCATTCAGGACCAAGTTATGTGGAACTGTTGTGTTTTCACAGAACAGCTATC 1985
Db 2111 GACGGCATTCAGGACCAAGTTATGTGGAACTGTTGTGTTTAAACAGAACGCTTAA 2170
Qy 1986 TATGTTATGAGCCCGCAATTAAGGCGAAGAGCGAGGTTTCTTGGCATCACTATGTTGG 2045
Db 2171 TATGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTTGGCTTGTGGG 2230
Qy 2046 GCAAGAAGAGCAAGCAAGTCAAGAAAAAGGAGCTCAGATAAGAAAAAGTCGAAACAAG 2105
Db 2231 GGAAGAAAAAGACGTCAAAATCTAAGAA--GAGCTCGAAAAAGAAAGTCACTAGAA 2287
Qy 2106 CATGTGACAGTCTCTGTTCCAGTATTTCAATCTCGAAGACATAGAGGAGGCTGTTAAGGT 2165
Db 2288 CACGACAGAGTCTCTGACAGTATTTAATCTCGAAGATATAGAGGAGGATGGAAGT 2347
Qy 2166 GCTGGGTTGATGATGAGAAATCAGTCTCATGTCTCAATGAGCTTAGAGAGAGATTT 2225
Db 2348 TCTCAGTTTGTATGAGAAATCGCTGATATGTCTCAATGAGCTTGGAGAGAGATTT 2407
Qy 2226 GCGCAGTCAGCAGCATTTGTTGCTCCACTCTCATGGAATATGTTGGTGTCTCTCAGTCC 2285
Db 2408 GGCAGTCCAGTGTGTTTGTAGCTCTACTCTGATGGAATATGTTGGTGTTCACATCT 2467
Qy 2286 TCCACTCCAGAAATCTCTTTTGAAGAGCTATCCATGTCTAAGTTGTGGCTATGAGGAC 2345
Db 2468 GCAACTCCAGAGTCTCTTCTGAAAGAGAGCTATTCATGTCTCAGCTGTGGCTATGAGGAC 2527
Qy 2346 AAGTCTGAATGGGAACTGAGATTTGGTGTGATCTATGATCTCTCAGAGAGATATCTT 2405
Db 2528 AAAAAGTGAAGGAACTGAGATTTGGTGTGATCTATGTTGCTTCTGTTACAGAGACATTTCTC 2587

QY 2406 ACTGATTCAAGATGACGACGAGGCTGGCTTTCAGTCTATTGTCATGCCAAGCGCCA 2465
|||
Db 2588 ACCGATTCAAGATGATGCTCGAGGCTGGCGATCAATCTACTGTCATGCTTAAGGACCA 2647
|||
QY 2466 GCTTTCAAGGGATCTGCCCGCCCAATCAATCTTTCAGATCGTGAACCAAGTGTGGGTGG 2525
|||
Db 2648 GCTTTCAAGGGATCTGCTCCCTATCAACCTTTCGGATCGTTTGAATCAAGTGTGGGTGG 2707
|||
QY 2526 GCTCTGGTCTGTGAAATCTTTTCAGCGCGCATTTGCCCTTATGATGATGCTTACGGA 2585
|||
Db 2708 GCTCTGTGTTCAATGAAATCTTTTCAGCAGGCATTTGCCATATGATGCTTATGGA 2767
|||
QY 2586 GGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCAACCATTTACCCATAC 2645
|||
Db 2768 GGGCGGCTAAATCTCTGAGAGATTTGCTTATATCAACACCAATTTATCCATCACA 2827
|||
QY 2646 TCTCTCCGCTTCTAGTCTATTGATATTGCTGCTATCTGCTCTCACTGGAAGTTC 2705
|||
Db 2828 TCAATCCGCTCTCTGCTACTGCTATTTGCGAGATTTGCTTCTCACTGGAGTTC 2887
|||
QY 2706 ATCATGCCAGATATGACACTTGGCCAGTATCTGGTTCATTTGCGCTCTTCTTCAAT 2765
|||
Db 2888 ATCATCCCAAGATATGAACTAGAGATGTTTGGTTTATATGCTCTTTATCTCAAT 2947
|||
QY 2766 TTCGCCACTGTTATCTTGAGATGAGTGGAGTGTGTCGATTCGACATCTGTTGCGCTCTTTCAGG 2825
|||
Db 2948 TTGCGCACTGGTATCTTGAGATGAGTGGAGTGTGTCGATTCGACATCTGTTGCGCTCTTTCAGG 3007
|||
QY 2826 AATGAACAGTTCTGGGTCAATTTGGAGTATCTCTGCACATCTGTTTGGCTCTTTCAGG 2885
|||
Db 3008 AACGAGCAGTTCTGGTCAATTTGGTATTTCTGCGCATTTATTTGCGCTCTTTCAGG 3067
|||
QY 2886 CTTCTGAAGGTCTTGGCGGTATTCGACCAACTTTCATCTGTCACCTCAAGGCTTAATGAC 2945
|||
Db 3068 CTCCTGAAGGTCTTGGTGTATTCGACAGAGCTTCACTGTCAACCTCAAGGCGCACTGAC 3127
|||
QY 2946 GAAGAGCGACTTGTGCTGAGTCTACATGTCAGTGGAGCGAGCTTCTATCCCTCG 3005
|||
Db 3128 GAAGAGGTGATTTTGGCGAGCTTACAGTTTCAAGTGGAGCAACGCTTCTGATCCCA 3187
|||
QY 3006 ACGACCAATTTTATCAATTAACATGTTGGTGTGCTGTCGACCTCTCTAGCCATCAAC 3065
|||
Db 3188 ACCACTATTTTATCATCAACCTGCTGGCGTGTGCTGGCATTTCTACGCATCAAT 3247
|||
QY 3066 AGTGTACCAATATGCGGCGCTCTTTGGGAAGCTTCTTCTTGGCTCTTGGGTGAT 3125
|||
Db 3248 AGCGTTTACCAATGATGCGGACCTCTTTTCGGAAGCTTCTTCTTGGGTGAT 3307
|||
QY 3126 GTTCACTTATACCAATCTCAAGGCTCTTATGGGAGGCAAAACCGCACCGACGAT 3185
|||
Db 3308 GTCACCTGTACCCCTTCTCAAGGCTCTATGGGAAGCAAGACCGCACCGGACCAAT 3367
|||
QY 3186 GTCATGCTGGGTGCTCTCTGCTTCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3245
|||
Db 3368 GTGTTGCTGGGTATCTCTCTGCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCT 3427
|||
QY 3246 CCATTCATACCGTCTGCTGGGCCCAATATCCAACTGTGGCATCACTGTGAGGA 3304
|||
Db 3428 CCATTCACACCGGCTCACTGGCCCTGATATCGGAAATGTGGCATCACTGTGAGGA 3486
|||

RESULT 10
AY055724
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY055724
Populus tremuloides cellulose synthase (Cesa5) mRNA, complete cds.
AY055724
AY055724.2 GI:39933009
Populus tremuloides (quaking aspen)
Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source

1. 3532
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
1. 3532
/gene="Cesa5"
/note="PtrCesa5"
236..3472
/gene="Cesa5"
/codon_start=1
/product="cellulose synthase"
/protein_id="PAL23710.2"
/db_xref="GI:39933010"
/translation="MESEGTGAKPMKSTGGVCOI CGDNVGTADGPFVACDVCAF
PVCPCYEYERKQDNQCPQCKTRYKLGSPAILGDEEDGADGASDFYSENQ
QKQIAERMSQWMTYGPEDGASNYDKEVSHNHIPLITNGHVSVDLSAEPHIS
MASPGAGGKRI PVASDVHSSNVVDPVPEFGSPGLGNVAKERVDMKMDKTV
VPMSTGHAPERGAGDIDAADVLVDLSDLEARQPLSRKVSIPSSKINPYRWIVL
RLGILCLFLHRIITNVPFNAYALGILYGTGWDFAISRLIDQFPKPNRETYDRL
ALRYDMGEPESHLVVDIFARSGVHLKEPPLVPANAVLSILAGSDPIDKVCYVEDDG
AAMUTFALSSETSSEFRKWFCKYISIEPAPWYFAQKIDYLDKQKVPQFVKDRRA
MKREYEFKIRINGLVAQKVPREGIMQDTEWPGNNTDHFQMIQVFLGSGGLD
SDGNELPRLVVSREKRPFOHKKAGANSLVRSVAVLTNGPFLNLDGCHYNISK
ALREAMCFMKNLGHVCYVQFPQDFGIDRNDRYANRNTVFFDINLRGLDGTGPV
YVGTGCVFNRTALYGPPLPKPKKPKGMLSSLCGSKSKSKSKSKSKSKGHV
DPTVPTSLDDIEBVGAGDDEKSLMSQMSLEKRPQSAFVASTLMENGVPQS
ATPTLLKEAHIHVSCEYEDKTDWSEITGTYGVEDILTFGKMHARGMSIYMPK
RPAFGKAPINLSDLNQLRWALGSVEILLSRHCPWYGGKRLKWLERFAYVNTI
YPTAIPLMYCTPAICLLTKFIIPOISNIAIWFISLFLSIPATGILBMRWSGVG
IDEMWNEQFWIGVSAHLPAVFGLLKVLADITNFTVTSKASDEGDSAEYLFX
WTLILPPTTLLIIVNLGVVAGISHAINSQSWGPLEGKLPFAPWIVHLYPLKGL
MGRQNRPTTIVVWMSILLASIFSLMLWRVDFTRVTPDVEQCGINC"

Query Match
Best Local Similarity 47.4%; Score 1720; DB 8; Length 3532;
Matches 2332; Conservative 0; Mismatches 875; Indels 27; Gaps 6;

rosids; eurousids I; Malpighiales; Salicaceae; Salicaceae; Populus.
1 (bases 1 to 3532)
Joshi,C.P. and Kalluri,U.C.
Isolation of two full-length Cesa genes from aspen trees homologous
to primary cell wall related Cesa genes from Arabidopsis
Unpublished
2 (bases 1 to 3532)
Joshi,C.P., Chavli,R., Kalluri,U.C. and Samuga,A.
Hypervariable II Regions from New Members of Aspen Cellulose
Synthase Superfamily Aid in Clarifying their Phylogenetic
Relationships with Other Cellulose Synthases in plants
Unpublished
3 (bases 1 to 3532)
Joshi,C.P. and Chavli,R.
Direct Submission
Submitted (11-SEP-2001) Forestry, Michigan Tech. University, 1400
Townsend Drive, Houghton, MI 49931, USA
4 (bases 1 to 3532)
Joshi,C.P. and Kalluri,U.C.
Direct Submission
Submitted (10-OCT-2002) Forestry, Michigan Tech. University, 1400
Townsend Drive, Houghton, MI 49931, USA
Sequence update by submitter
On Dec 16, 2003 this sequence version replaced gi:23428423.
Location/Qualifiers
1. 3532
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
1. 3532
/gene="Cesa5"
/note="PtrCesa5"
236..3472
/gene="Cesa5"
/codon_start=1
/product="cellulose synthase"
/protein_id="PAL23710.2"
/db_xref="GI:39933010"
/translation="MESEGTGAKPMKSTGGVCOI CGDNVGTADGPFVACDVCAF
PVCPCYEYERKQDNQCPQCKTRYKLGSPAILGDEEDGADGASDFYSENQ
QKQIAERMSQWMTYGPEDGASNYDKEVSHNHIPLITNGHVSVDLSAEPHIS
MASPGAGGKRI PVASDVHSSNVVDPVPEFGSPGLGNVAKERVDMKMDKTV
VPMSTGHAPERGAGDIDAADVLVDLSDLEARQPLSRKVSIPSSKINPYRWIVL
RLGILCLFLHRIITNVPFNAYALGILYGTGWDFAISRLIDQFPKPNRETYDRL
ALRYDMGEPESHLVVDIFARSGVHLKEPPLVPANAVLSILAGSDPIDKVCYVEDDG
AAMUTFALSSETSSEFRKWFCKYISIEPAPWYFAQKIDYLDKQKVPQFVKDRRA
MKREYEFKIRINGLVAQKVPREGIMQDTEWPGNNTDHFQMIQVFLGSGGLD
SDGNELPRLVVSREKRPFOHKKAGANSLVRSVAVLTNGPFLNLDGCHYNISK
ALREAMCFMKNLGHVCYVQFPQDFGIDRNDRYANRNTVFFDINLRGLDGTGPV
YVGTGCVFNRTALYGPPLPKPKKPKGMLSSLCGSKSKSKSKSKSKSKGHV
DPTVPTSLDDIEBVGAGDDEKSLMSQMSLEKRPQSAFVASTLMENGVPQS
ATPTLLKEAHIHVSCEYEDKTDWSEITGTYGVEDILTFGKMHARGMSIYMPK
RPAFGKAPINLSDLNQLRWALGSVEILLSRHCPWYGGKRLKWLERFAYVNTI
YPTAIPLMYCTPAICLLTKFIIPOISNIAIWFISLFLSIPATGILBMRWSGVG
IDEMWNEQFWIGVSAHLPAVFGLLKVLADITNFTVTSKASDEGDSAEYLFX
WTLILPPTTLLIIVNLGVVAGISHAINSQSWGPLEGKLPFAPWIVHLYPLKGL
MGRQNRPTTIVVWMSILLASIFSLMLWRVDFTRVTPDVEQCGINC"

QY 83 GAAGTCGGGAGGACCGGGCGGGGAGCTGTGCGAGATCTGCGCCGACGCGCTGGGAC 142
|||
Db 262 GAAGCCATGAAAGTACAGGTGGACAGTTTGGCAGATCTGTGTGATATGTTGCAA 321
|||
QY 143 CACGTTGACGCGACGCTTTCACCGCTGCGAGCTGCGCGCTTCCGCGTCTCGCGCC 202
|||
Db 322 GACAGCGATGTTGAACCGTTTGTGTTGCGATGTTTGTGCAATTCAGTGTGCGGCC 381
|||
QY 203 CTGCTACGACGACGCGAGGAGGACCCGCGCTGCTCCAGTCGACAGCAAGTA 262
|||

Db 382 TTGCTATGAGTACGAGGAGGATGGCAATCAGTCTTGCCCTCAGTGCAGAAACACAGATA 441
Qy 263 CAAGCCGACAGAGGAGAGCCAGCGATCCCGGGGAGGAGGACGACACTGATGCCGA 322
Db 442 CAAGAGGCTCAAGAGGTAGCCCTGCCATCTTGGTGATAGAGAAAGGATGGTGATCTGA 501
Qy 323 TGATGCTAGTGACTTCAACTACCTGATCTGATCTGAGGACGAGAGCAGAAAGATTCG 382
Db 502 TGATG--GTGCTAGTGAATTTAACTTACTCAGAAATCAAAACAGAAACAGAAAGATTC 558
Qy 383 TGACAGATGCGCAGCTGGCGCATGAACACCGGGGGCAGTGGCAATGTTGGCCACCCCAA 442
Db 559 GGAACGATCTGAGCTGGCAGATGACTTATGACACAGGGGAAAGATTTCTGGAGCTTCAA 618
Qy 443 GATGACAGTGGCGAGATCGGCCCTCTCAAGTATGACAGTGGAGAGATCCCTAGGGGATA 502
Db 619 TTATGATAAAGAGG---TTTCGCATAAACCATATCCACTACTCACCAAATGGACATGAGG 674
Qy 503 CGTCCCTTCAGTCACCAACAGCCAGATGTGAGGAGAAATCCCTGGAGCTTCGCTGATCA 562
Db 675 TTCTGGAGATCTGTCTGCTGATCAGCTGAGCATATTTCTATGGCTCTCTCGGTGCTG 734
Qy 563 TCACATGATGTCCTTACCGGGAACATCAGCAGACGTGCTCGTTCCTTATGTGAATCA 622
Db 735 GTGGTGGGAAGCGATCCCTTATGCACTCTGATGTCATCAATCATCTAATGTGAGAGTTG 794
Qy 623 TTCACCAATCCGTCAAGGAGTTCTCGGCAGTATTTGGAAATGTTGCCCTGGAAAGAG 682
Db 795 TGGACCCAGTAAGGAGTTTGGTTTACCGGGACTG--GGCAATGTGGCCCTGGAAAGAAAG 852
Qy 683 AGTTGATGGCTGAAATGAAGCAGGACCAAGGTGGATTCCTCATGACTAATGGGACAAG 742
Db 853 AGTTGATGGCTGAAAGATGAAGCAGGACAAGACTGTTTCTTATGAGTACTCG----- 906
Qy 743 CATTTGCTCCCTTGAAGGTGGGAGCTACTGATCGATGCACTACTGCAATACAAAT 802
Db 907---TCATGCTCCTTCTGAAAGGGGTGCTGGAGATATTTGATGCTCTACGGATGTGCTGT 963
Qy 803 GGAAGAGCTTTACTGATGATGAACCTCCGAGCCTCTATCTAGAAAGTCCCGATTGC 862
Db 964 GGAATGCTTTTACTGATGATGAAGCTCGCAACCTCTTTCGAGGAAGTTTCCATPCC 1023
Qy 863 TTCTCTCAAAATAAATCCCTACAGATGGTCAATTTCTCGCGTGGTGTCTTCTAAGCAT 922
Db 1024 TTCTTTCTAAGATTAATCCATACAGATGGTATTTGTTTTCGGCTTGGTATCTCTGCA 1083
Qy 923 CTTCCTGCACTACCGTCTCAAAATCTGTGCGTAAATGCAATCCCACTGTGGCTTTATC 982
Db 1084 TTCTTTGCACTACCGAATAACAAATCTGTGCGTAAATGCGTATGCTCTGGGGTTATATCT 1143
Qy 983 TGTATATGTGAGATTTGGTTTGTCTTATCTCGATCTGATCAGTTCCCGAAGTGTT 1042
Db 1144 GGGGATATGGGAGATTTGGTTTGGATATCCCGGATCTGATCAGTTCCCTAAGTGGCT 1203
Qy 1043 TCCAATCAACCGGGAGACCTACCTTGATGACTGGCTTTAAGGTATGACCGAGAGGTGA 1102
Db 1204 CCCTGGAAATCGGGAACATATCTTGACAGCTTGTCTTGAGATATGACATGGAAGAGA 1263
Qy 1103 ACCGTCTAGTTGGCTGCTGTGACATATTTGTGCTAGTACAGTCAAGTCCCTTGAAGAGCC 1162
Db 1264 ACCCTCCACTTGGTGGTGGACATTTTGGCAGGTCCGGGTCTCCATTTAAAGAGGCC 1323
Qy 1163 ACCTATGCTCACTGCCAACHCTGTCTATCCATTTCTGCTGTTGATTTATCCCTGGACAA 1222
Db 1324 CCCTTTGGTCCAGCAATGAGTGTCTATTTCTTTCGAGGTGACTCCCAATTTGACAA 1383
Qy 1223 GGTCTCTTGTATGTATCTGATGACGAGCTTCAATGCTGACTTTTGGACGATTTGGCTGA 1282
Db 1384 AGTCTCATGCTATGCTCTGATGATGGTGTGCAATGTTGACATTTGAAGCTTTGTCTGA 1443
Qy 1283 GACTTCAGATTTGCTAGGAAATGGGTACATTTGTGAGAAAGTATGACATTTGAACCCAG 1342
Db 1444 GACTTCAGAGTTTTCAGGAAATGGGTGCTTCTGCAAGAGTATGACATTTGAGCCTCG 1503

Qy 1343 AGCTCCGAGTCTTTACTTTTCCAGAAAATTGATTAACCTGAAAGACAAAGTCCAGCTTC 1402
Db 1504 GGCTCCAGAAATGGTACTTTTGCAAAAAGATTTGACTACTTGAGGATTAAGTGCAGCAATC 1563
Qy 1403 ATTGTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAGAAATTTAAATACGAGATAA 1462
Db 1564 TTTTGTCAAAGACCGTAGAGCTATGAAGAGAGAAATATGAAGAAATTTAAATTCGTATCAA 1623
Qy 1463 TGGCCCTAGTTTCTAAGGCATTTGAAAGTCCCGAGGAGAGATGGATCATGCAAGATGGCAC 1522
Db 1624 TGGGCTTGTGTGCAAGCGCAAAAGGTGCTGGAAGAGGATGGATTTATCAAGATGGGAC 1683
Qy 1523 ACCATGGCCAGGAAACAATACCAGGGATCATCTCGAAATGATTCAGGTTTCTCTGGTCA 1582
Db 1684 CCCATGGCTTGGAAATTAACACAGAGACCATCAGGAATGATTCAGGTTTTCTTGGGCCA 1743
Qy 1583 CAGTGTGGCTTTGATPACTGAGGGTAAATGAGCTCCCGCTTTAGTTTATGTGTCTCGTGA 1642
Db 1744 AAGTGGAGGGCTTGAATCTGATGTTAATGAGCTTCCGCGTTAGTTTATGTTTCTCGTGA 1803
Qy 1643 AAGCGTCTTGGTTCAGCACCAAGAGGCTGTGCCATGAATGCCCTTGTCTGTGT 1702
Db 1804 AAGCGTCTTGGCTTCCAGCACCAAGAGGCTGTGCCATGAATTCACCTTGTCTGTGT 1863
Qy 1703 CTCAGCTGTCTTACTAATGGCAATACATGTTGAATCTTGATTTGTGATCAGTACATCAA 1762
Db 1864 ATCAGCGTGTCTAATCTAATGGGCTTCTTGTGAAATCTTGATTTGATCAGTACATAAA 1923
Qy 1763 CAACAGCAAGGCTGTCCGAGAGCTATGTCTTCTTAATGGATCCAAAACCTAGTCCGCA 1822
Db 1924 CATTTAGCAAGGCTTTGAGAGAGCTATGTGTTTTCATGAAGGATCCCTAACCTTGGGAAACA 1983
Qy 1823 AGTCTGTTATGTCAGTTCCCAAGAGGTTTGTGAGATTTGATAGGAATGATCGATATGC 1982
Db 1984 TGTCTGTATGTTCAGTTTCCAAAGATTTGATGCAATTCAGAGGAATGATCGATATGC 2043
Qy 1883 AAACAGGAACACTGTCTTTTTCATATTAACCTTGAGGGCTTTGACGGCAATTCAGAGACC 1942
Db 2044 CAATCGCAATACGTATTTCTTTGATATAAATCTTGAGGGGTCTGAGTGTATTTCAAGGACC 2103
Qy 1943 AGTTTATGTGGAACTGGTTGTGTTTTCAACAGAAACAGCTATCTATGTTTATGAGCCCC 2002
Db 2104 TGTCTATGTGGGTACTGGATGTCTTTAAACAGACGGCATGTATGTTTATGAACTCC 2163
Qy 2003-----AATTAAGCCAGAGAGCGAGTTTCTTGGCATCACTATGTGGGG---CAAGAA 2053
Db 2164 CCTCAACCCAGCATTAAGAAACAGGCAATGCTCTTCACTCTGTGGTGGATCAAGAAA 2223
Qy 2054 GAAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAAAGTCCGAAACAAGCATGTGA 2113
Db 2224 GAAGGGTTCTAATCAAGCAAAAAGGGAATCTGACAGAGAAATCTGGAAGCATGTAGA 2283
Qy 2114 CAGTTCTGTTCCAGTATTCAAATCTCGAAGCATAGAGGAGGGGTGTTGAAGGTGCTGGGTT 2173
Db 2284 TCCCACTGTACCCATTTTCAGCTTGGACGATATAGAAAGGGCTTTGAAGGAGCTGGATT 2343
Qy 2174 TGAATGATGAATCAGTTCTCATGTCTCAAAATGAGCTTAGAGAAGATTTGCGCCAGTC 2233
Db 2344 TGAATGAGAGAGTCACTTCTCATGTCTCAAAATGAGCTTGGAGAGAGGTTTGGCCAGTC 2403
Qy 2234 AGCAGCATTTGTTGCTCCACTCTGATGGAATATGTTGGTGTTCCTCAGTCCCTCCACTCC 2293
Db 2404 AGCTGTTTTTGTGCTTCCACACTTATGGAGATGGTGGTGTTCCTCAGTCTGCAACACC 2463
Qy 2294 AGAATCTCTTTGAAAGAGCTATCCATGTATGATTAAGTTGTGGCTATGAGGACAAAGTCTGA 2353
Db 2464 AGAACTCTTCTTAAAGAGCTATTCATGTCTATGTTGTGGGTATGAGGACAAACAGA 2523
Qy 2354 ATCGGGAATGAGATTTGGTGTGATCTATGATCTGTACAGAAAGATTTCTTACTGGATT 2413
Db 2524 TTGGGATCTGAGATTTGGATGGAATCTATGTTTCTGTTACAGAAAGATTTCTTACTGGGTT 2583

QY 2414 CAAGATGACGACGAGAGGTGGCGTTCACTATTGTCATGCCCAAGCGCCAGCTTTCAA 2473
 DB |||||
 QY 2584 TAAAGATGATGCCGGGGTGGCGGTCAATTATTGTATGCCAAGCGCCAGCCTTTAA 2643
 DB |||||
 QY 2474 GGGATCTGCCCCCAATCAATCTTTTCAGATCGTCTGAACCAAGTGTGGGTGGGCTCTCGG 2533
 DB |||||
 QY 2644 AGGGTCTGCTCTAATTAATCTTTCTGATCGTCTGAACCAAGTGTTCGATGGCTCTGGG 2703
 DB |||||
 QY 2534 TTCTCTTGAATCTTTTCAGCGCGGATGCGCCCTTATGGTATGGTACGAGGGGGCT 2593
 DB |||||
 QY 2704 TTCTCTTGAATCTTTTCAGCGCGGATGCGCCCTTATGGTATGGTATGGTATGGT 2763
 DB |||||
 QY 2594 CAAGTCTCTGGAGATTCGCTTACATCAACACCACTTTACCCCACTAACTCTCTCC 2653
 DB |||||
 QY 2764 GAAATGGCTCGAGAGTTGCAATATGTAACACTACCATTTATCATCACCGCATTC 2823
 DB |||||
 QY 2654 GCTTCTAGTCTATGATATTCCTGCTATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 2713
 DB |||||
 QY 2824 TCTTCTCATGATTCGACACTTCCAGCTATCTGCTACTTACAGACAAATTCATTATTC 2883
 DB |||||
 QY 2714 AGAGATTAGCACTTTGCGCAGTATCTGTTCAATGGCTCTCTCTTCAATTTTCCAC 2773
 DB |||||
 QY 2884 CCAGATTAGTACATCGCAAGCATATGGTTATTTCTCTCTTCTCTCTCTCTCTCT 2943
 DB |||||
 QY 2774 TGGTATCTCTGAGATGAGTGGTGTGTCGCTTACAGATGCTGGAGGAAATGAACA 2833
 DB |||||
 QY 2944 CGGTATCTTGGAGATGAGTGGTGTGTCGCTTACAGATGCTGGAGGAAATGAACA 3003
 DB |||||
 QY 2834 GTTCTGGGTCAATGAGGATCTCTGCAATCTGTTGCGCTTTCAGGGTCTTCTGAA 2893
 DB |||||
 QY 3004 GTTCTGGGTCAATGAGGATCTCTGCAATCTGTTGCGCTTTCAGGGTCTTCTGAA 3063
 DB |||||
 QY 2894 GGTGCTTCCGCTATCGACCAACTTCACTGCTCACTCAAGGCTATGACGAAGAAG 2953
 DB |||||
 QY 3064 AGTACTTCTGTCATGACCAACTTCACTGCTCACTCAAGGCTATGACGAAGAAG 3123
 DB |||||
 QY 2954 CGACTTCTGTCAGCTCTACATCTCAAGTGGAGAGCTTCTCATCCCTCCGACCACT 3013
 DB |||||
 QY 3124 TGATCTGCTGACTATCTGTTCAATGGACACTTACTTATCCCGCCCAACTCT 3183
 DB |||||
 QY 3014 TTTGATCAATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3073
 DB |||||
 QY 3184 CTTATCTGTTAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3243
 DB |||||
 QY 3074 CCAATCATGGGGCGCTCTTTGGAGAGCTCTTCTTGTGCTTCTGGGTGATTGTTCACT 3133
 DB |||||
 QY 3244 TCAGTCATGGGGTCCACTCTTTGGCAAGCTTCTTCTGCTTCTGGGTGATTGTTCACT 3303
 DB |||||
 QY 3134 ATACCCATTCTCCCAAGGCTTATGGGAGGCAAAACCGCACACGACGATTGTCATCT 3193
 DB |||||
 QY 3304 CTATCTTCTTAAAGTTTGTATGGTCCGAGAACCGGACACCACTTGTGTTGTT 3363
 DB |||||
 QY 3194 CTGGGCT 3253
 DB |||||
 QY 3364 GTGGTCTATCT 3423
 DB |||||
 QY 3254 TACCCGCTCTGCTGGCCCAATATCAACCTGCTGGCATCACTGCTAGGAAG 3307
 DB |||||
 QY 3424 CATTAGAGTTACTGTTCCAGATGTTGAGCAGTGGCGGCACTCACTGTTAGAGAAG 3477
 DB |||||

RESULT 11

AF150630
 LOCUS Gossypium hirsutum 3723 bp mRNA linear PLN 04-FEB-2000
 DEFINITION Gossypium hirsutum cellulose synthase catalytic subunit (celA3)
 mRNA, complete cds.
 ACCESSION AF150630 AF200453
 VERSION AF150630.2 GI:6446576
 KEYWORDS
 SOURCE Gossypium hirsutum (upland cotton)
 ORGANISM Gossypium hirsutum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS
TITLEJOURNAL
AUTHORS
TITLEJOURNAL
AUTHORS
TITLEJOURNAL
AUTHORS
TITLEJOURNAL
AUTHORS
TITLEREMARK
COMMENT
FEATURES

source

gene

CDS

ORIGIN

Query Match 46.9%; Score 1699.8; DB 8; Length 3723;
 Best Local Similarity 72.6%; Pred. No. 0;
 Matches 2346; Conservative 0; Mismatches 827; Indels 60; Gaps 9;

QY 84 AAGTCCGGAGGACGCGCGCGGACGCTGTGCCAGATCTGCGCGGACGCGCTGGGAC 143

DB 296 AAGCCCATGAAGAATCTTTGGTGGCAGACATGCCAAATCTGTGGTGAATGTTGGCAA 355

QY 144 ACGTTGAGCGGACGCTTTCACGCGCTGCGACGCTCTCCGCTTCCCGGCTTCCCGCCCC 203

DB 356 AATACAGATGTTGATCCATTTCATGCTTGCATATATGTCATTCCTGATGAGGCT 415

rosids: eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 3723)
 Laosinchai, W., Cui, X. and Brown, R.M. Jr.
 A full length cDNA of cotton cellulose synthase has high homology
 with the Arabidopsis RSM1 gene and the cotton CelA1 gene (Accession
 No. AF200453) (PCR 00-002)
 Plant Physiol. 122 (1), 291 (2000)

2 (bases 1 to 3723)
 Kimura, S., Laosinchai, W., Itoh, T., Cui, X. and Brown, R.M. Jr.
 Immunogold Labeling of Rosette Terminal Cellulose Synthesizing
 Complexes in a Vascular Plant (Vigna angularis)
 Unpublished

3 (bases 1366 to 3470)
 Laosinchai, W. and Brown, R.M. Jr.
 Direct Submission
 Submitted (11-MAY-1999) Botany, University of Texas at Austin, W.
 24th Street, Austin, TX 78713, USA

4 (bases 1 to 3723)
 Laosinchai, W., Cui, X. and Brown, R.M. Jr.
 Direct Submission
 Submitted (17-NOV-1999) Molecular Genetics and Microbiology,
 University of Texas at Austin, Austin, TX 78712, USA

Sequence update by submitter
 On Nov 17, 1999 this sequence version replaced gi:5081778.
 Location/Qualifiers
 1..3723

/organism="Gossypium hirsutum"

/mol_type="mRNA"

/cultivar="Texas marker-1"

/db_xref="taxon:3635"

/tissue_type="cotton fiber"

/dev_stage="24 days post anthesis"

1..3723

/gene="celA3"

269..3472

/gene="celA3"

/note="celA1 homolog"

/codon_start=1

/product="cellulose synthase catalytic subunit"

/protein_id="AAD39534.2"

/db_xref="GI:6446577"

/translation="WESGDIIGKPMKNIIGGTCQICGDNVGNKNTDGPFIACNICAF
 PVCPCYEYERKDNQSCPOCKTRYKWKQGPSAILGDRGTGGDADGASDFIYSENQ
 QKQLAKRMQWNAKIRGVEDGAPTYDEIKSHNHIPILTSQGVSGVSLASAPRLS
 MASPGVAGKGSIRVDPBFGSLGNVAKMVERKVDGKMKQKNTVPMSTCQATSE
 RGLGDIDASTDVLVDQNDDEARQLSRKVSQSSKINPYRMVILRLVILICILHY
 RITVPNAYALWLTIVICETFAISWILDQPKWLPVNRETYDLRLAIVDRGEPHS
 ELAAVDIFVSTVDPLKBPPLVANTVLSILAVDPVKVSCVSDGGAAMLTFEALSE
 TSEFARKWVPECKXNIEPRAPWEVQAKIDYLDKQVTSFVKORRAMREYEFKVR
 INGLVAKAQKVPBEGIMQDTPWGNTRDHPGMIQVFGSGGDAEGNELPLRVY
 VSRKRPQFQHKKAGAMNALVRVSALITNGAFLLNLDCHYINNKKALREACFLMD
 PNLQKQCYVQFPQRFQDIDRNDYRANTVFFDINLRGLDGIQGFVVTGTCVFNRT
 ALYGEVPLKPKHRTKTLISLSCGSRKSKSSKSGSKGKHVDSTVPVFNLED
 IEVGEVAGFDDKSLMSOMSLKRFQGSVFAVSTLMWNGVQPSATPFTLLKEAI
 HVIQSGYEDKTDWGEIGWIVGVTEDILTGFKMHARCWRSTYCMQPKPAKGSAPIN
 LSRDLNQLVRNALSVELLSRHPCIWYSGRLKWLERFAYVNTIYPTVAILPMY
 CTLFVACLTNTKFIPIQISNLASIFLSLFIATGILKRWKNGVTDQWNRNQFW
 VIGGVAHLFAVFGQLKLVGIDITNFTVTSKASDEGDFAEILYMKFTLLIIPPTLL
 LIINLVGVAGISYVINGSGWGLFGKLFQFAFWVIIHLYPFLKGLMGRQNRTPTV
 VVMSILLASIFSLWVRIDPPTTRVTRVTPDVEQCGINC"

[illegible]

ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
AUTHORS	1 (bases 1 to 3229) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.			
TITLE	Arabidopsis Open Reading Frame (ORF) Clones			
JOURNAL	Unpublished			
AUTHORS	2 (bases 1 to 3229) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-DEC-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA			
COMMENT	Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.			
FEATURES	Location/Qualifiers			
source	1..3229 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="5" /clone="C104938" /ecotype="Columbia" /note="This clone is in pUNI 51."			
gene	1..3229 /gene="At5g05170"			
CDS	1..3198 /gene="At5g05170" /codon_start=1 /evidence=experimental /product="putative cellulose synthase catalytic subunit" /protein_id="AA86168.1" /db_xref="GI:26983832" /translation="MSEBETAGKPMKNIVPQTQICSDNVGKTVDPDRFVACDICSFVPCPCVEYRKDGNQSCPOCKTRYKRLGSPAIQGDQEDGLADGETVEFNYPQKEKISRMKLHTRCKGEMGEPOYKRVSHNHLPLTSDRTSGEFSAAAPSVLSVSTIAGKRLPYSSDYNQSPNRRIYDPVGLGNVAKERVGDGKWKQKERTKGPVSTCAASERGVDIADATL1DAEALLDEARQPLSRKVSIPSSRNPNRYMRLRLVILCLFLHYRITNPENAFALWLVSCICEIWFALSWILDQFPKFPVNRETYLDRLALRYDREGESPSQLEARDIYFVPTCKEPLPVANTVLISILADYVDVKVSCYSDSDGAAMLPSLEALRISFAVQKVDKSWGMEIQTWRAPWEYFAAKIDYLDKQVTSFKDRAKMRKREYEFKIRIISVXKALCPKEGVQMDQTPWGNTRDHPGMIQVFLGONGGLDAEGNELPLVVVSRKRGFQHQHKKAGAMALRVYSAVLTNGPFTILNDGDHYINNSKALRAMCFLMDP NLKQVCYVQFPQRFQGLKNDRYANRNVFFDINLRGLDGIQGPVYVGTGCVENRTA LYGEPIPKVHKHKSULLKCGSRKKSNAKESDKKSGRHTDSTVDFVNLDDIEVEGEVGFDDKALLMSQLEKRFQCSAVFVASTLMNGVPPSPATPENLKAELRISCSQYEDKDWGMEIQTWRAPWEYFAAKIDYLDKQVTSFKDRAKMRKREYEFKIRI DRNQVRLWALGSVEILFSRHCPITWYNGRLKFLERFAVNTIYPTISIPUMKTSI LPVCLFTNQFIIPQISNLAS1WFLSFLSIFATGILEMWSGVGDWWRNEQFVVI GGVS4HLFAVGQILKVLKGLADTNFTVTSKASDEGDFAEIYFLKMTTLLIPPTILLI WSVLLASIFSLWLVRIDPFTSRVTSGPDILCECGNC" 3199..3229 /gene="At5g05170"			
3'UTR				
ORIGIN				
Query Match	46.3%; Score 1679.4; DB 8; Length 3229;			
Best local Similarity	72.0%; Pred. No. 0;			
Matches 2333; Conservative	0; Mismatches 836; Indels 72; Gaps 8;			
Qy	77	CGCCTGAAGTCGGAGGACAGGGCGGGACGTGTGCCAGATCTGCCGACCGCCT	136	
Dd	21	CGCGGAAACCCATGAAGAACATGTTCGCGACATTCGCAGATCTGTAGTCAATGT	80	
Ov	137	GGGCACACGTTGGACGGGACGCTTCTCACCGCCTGCAGCTTCGCCGCTTCGCCGCTC	196	

81 TGGCAAGACTGTGTGATGGAGATCGTTTGTGGCTTGTGATATTTGTTCAATCCAGTTTG 140
197 CGGCCCTGCTACGACGAGCGCAGAGGAGCAACCGAGGCTGCTCGCTCAGTGCAGAC 256
141 TCGGCCCTGCTACGAGTATGAGAGGAAAGATGGAAATCAATCTGTCTCAGTGCAGAAC 200
257 CAAGTACAGCGCCACAGAGGAGCCAGCGATCCGCGGGGAGGAGGACACACTGA 316
201 CAGATACAGAGGCTCAAGAGGTAGTCTGCTATCTCTGTGTGATAAAGACAGAGATGGCTT 260
317 TGGCGATGATGGTA---GTGACTTCACTACCTCGCATCTGGCACTGAGGACAGAGCA 373
261 AGCTGATGAGGTACTGTGTGAGTTCACCTACCT-----CAGAAGGA 302
374 GAAGATTGCTGACAGATGCGAGCTGGCGCATGAACACCGGGGGCAGTGGCAATGTGG 433
303 GAAAAATTCAGAGCGGATGCTGTGTTGGCATCTTACTCGTGGGAGGAGGAAATGGG 362
434 CCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAGTATGACAGTGGAGAGATCCC 493
363 GGAACCCAGTATGATAAGAGG-----TCTCTCAATCATCTTCTCTGCTCTCAC 413
494 TAGGGATACGTCCTCTTCACTACCAACAGCCAGATGTTCAGGAGAAATCCCTGGAGCTTC 553
414 GACGACACAGATACT-----TCAGGAGAGTTTTCTGCTGCCCTC 452
554 GCCTGATCATCATGATGTCCCTACGCGGGAACATCAGCAGAGCTGCTCCGTTCCCT- 612
453 ACCTGAACGCTCTCTGTATCTTCTACTATCGCTGGGGGAAAGCGCTTCCCTATTATC 512
613 --ATGTGAATCATTCACCAATCCGTCAAGGGAGTTCTCGGCAGTATTCGGGAATGTTC 670
513 AGATGTCATCATCAATCAAAATAGAGGATGTGGATCCTGTTGGACTCGGGAATGAGC 572
671 CTGGAAGAGAGAGTTGATGGCTGGAAATGAAGCAGGACAAAGGGTGCATTCCTCATGAC 730
573 TTGGAAGGAGAGAGTTGATGGCTGGAAATGAAGCAGGACAAATACTGGTCTCTGTCAG 632
731 TAATGGACAAGCATTTGCTCCTCTGAGGTGGGAGCTACTGACATCATGATCATAC 790
633 CAGCGAG-----GCTGCTCTGTAAGAGGTGGAG---TAGATATTGATGCCAGCAC 680
791 TGAATCAACATGAAGAGCCTTTACTGTAATGATGAACCTCGCCAGCTCTATCTAGAAA 850
681 AGATATCTAGCAGATGAGGCTCTGCTGAATGACAGCAGGAGGCTCTGTCAAGGAA 740
851 AGTCCCATTTGCTTCTCCAAAAATAATCCCTACAGAAATGGTCAATTTCTGCGGTTGGT 910
741 AGTTTCAATTTCTTCATCAGGATCAATCTTACAGAAATGGTTATTATGCTGCGGCTGT 800
911 TGTTCTAAGCATCTTCTGCACTACCGTCTCACAAAATCCTGTGGTAAATGCATACCCACT 970
801 TATCCTTTGCTCTTCTGTGATTAACCGTATAAACAACCCAGTCCAAATGCCCTTTGCTCT 860
971 GTGCTTTTATCTGTATATGTGAGATTTGGTTGCTTTATCTCTGGGATCTGGATCAGTT 1030
861 ATGCTGGTCTCTGTGATATGTGAGATCTGGTTGGCTTATCTCTGGATTTTGGATCAGTT 920
1031 CCCGAAGTGTTCATCAACCGGGAGACCTACCTTGTATGAGACTGGCTTTAAGGATGA 1090
921 TCCCAAGTGTTCCTGTGAACCGTGAACCTTACCGTATAAACAACCCAGTCCAAATGCCCTTTGCTCT 980
1091 CCAGAGAGTGAACCGCTCAGTTGGCTGTGTTGACATATTTGTCACTACAGTCCGACCC 1150
981 TCGTGAAGGTGAGCCATCAGAGTTAGTCTGCTGTGACATTTTCGTGAGTACTGTGTGACCC 1040
1151 CTTGAAGGAGCCACTTCTGTCAGTGCACACTGTGCTATCCATTTCTGTGTTGATTA 1210
1041 CTTGAAGGAGCCACCTTGTGACAGCAACACAGTCTCTCTATTTCTGCTGTGACTA 1100
1211 TCCCGTGGACAGGTCTCTGCTATGATGATGACGAGGACTTCAATGCTGATCTTTGA 1270

1101 CCCAGTTGACAAGGTGTCTCTGTTATGTTTCTGATGATGGTGTCTATGTTATCATTTGA 1160
1271 CGCATTTGCTGACACTTCAGAGTTTCTAGGAAATCGGTACCAATTTGTGAAGAGATGA 1330
1161 ATCACTTGAGAAACATCAGAGTTTCTGCTAAATGGGTACCAATTTTGAAGAAATATAG 1220
1331 CATTTGAACCCAGAGCTCCCGAGTTTACTTTTCCAGAAAAATTTGATTACCTGAGAAACAA 1390
1221 CATAGAGCTCTGTCACAGAAATGGTACTTTGCTGCGAAAAATAGATTACTTTGAAGGATA 1280
1391 AGTCCAGCTTCATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATAGAAGAAATTTAA 1450
1281 AGTTCAAGATCATTTTGTCAAGATCGTAGAGCTATGAAGAGGGAATATGAGGAAATTTAA 1340
1451 AATCAGATTAATGCGCTAGTTTCTAAGGCATTTGAAGTCCCGAGAGAGATGGATCAT 1510
1341 AATCCGAATCAATGCACTTTGTTTCCAAAGCCCTTAAATGTCTTGAAGAGGTTGGGTAT 1400
1511 GCAAGATGGCACACCATGCGCAGGAAACAATACAGGGGATCATCTTGGAAATGATTCAGGT 1570
1401 GCAAGATGGCACACCGTGGCTGGAATAATACAAGGACCATCCAGGAATGATCCAGGT 1460
1571 TTTCTTGTGTCACAGTGGTGGCTTGTATCTGAGGGTATGAGCTCCCGCTTTAGTTTA 1630
1461 CTTCTTAGGCAAAATGTTGAGCTTGTATGAGAGGCAATGAGCTCCCGCTTTGGTATA 1520
1631 TGTGCTCTCGTGAAGAGCTCTCGGTTCCAGCACCAACAAGAGGCTGGTCCCATGAATGC 1690
1521 TGTTTCTGAGAAAGAGGACAGGATTCAGCACCAACAAGAGGCTGGTCTATGAATGC 1580
1691 CTTTGTTCGTGCTCAGCTGTCTTACTAATAGCAATACATGTTGAATCTTGAATGTGA 1750
1581 ACTGTTGAGAGTTTTCAGCAGTTCTTACCAATGACCTTTTCTTGAATCTTGAATGTGA 1640
1751 TCACTACATCAACACAGCAGGCTGTCCGAGAGCTATGCTTCTTCTAATGATCCAAA 1810
1641 TCATTATATAATTAACAGCAAGCCTTTAAGAGAGCAATGCTTCTCTGATGAGCCAAA 1700
1811 CCTAGTCCGCAAGTCTGTTATGCTGCAAGTCCCAAAAGGTTTGTAGGGATTCATAGGAA 1870
1701 CCTCGGAAGCAAGTTTGTATGTTAGTTCAGTTCACAAAGATTTGATGATTCAGTAAGAA 1760
1871 TGATCGATATGCAACACAGGAACTGCTGCTTTTGTATATTAATCTGAGGGCTTGAAGG 1930
1761 CGATAGATATGTAATCGTAAATACCGGTGTTCTTTGATATTAATTAATCTGAGAGGTTTAGATGG 1820
1931 CATTCAGGACCAAGTTTATGTTGGAAGTGGTGTGTTTCAACAGAACACTATCTATGG 1990
1821 GATTCAAGGACCTGTATATGTCGGAACCTGATGTTGTTTCAACAGAACAGCAATATACGG 1880
1991 TTATGAGCCCCCAATTAAGG-----CGAAGAGCCAGGTTTCTTGGCATCATCTATGTTG 2044
1881 TTATGAACCTTCAATTAAGTAAACACAGAGAGCCAGTCTTTTATCTAAGCTCTGTTGG 1940
2045 GGGCAAGAGAGGCAAGCAAGTCAAGAAAAAGAGCTCAGATAAGAAAAAGTCAAGCAAA 2104
1941 TGGATCAAGAAAGAGAAATCCAAAGCTAAGAAAGAGTCCGACAAAAAGAAATCAGGAC 2000
2105 GCATGGAAGCTTCTGTTCCAGTATTCATCTGAGACATAGAGAGGCTGTTGAAGG 2164
2001 GCATCTGACTCAACTGTTCTCTGTTTCAACCTCGATGACATAGAGAGGAGTTGAAGG 2060
2165 TGCTGGGTTGATGATGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAGAGAGATTT 2224
2061 TGCTGTTTGTATGATGAAAGGCGCTCTTAATGTCGCAATGAGCTTGAGAGCCGATT 2120
2225 TGSCCAGTACAGAGCATTTGCTCCTCACTCTGATGGAATATGTTGGTGTTCCTCAGTC 2284
2121 TGGACAGTCTGCTGTTTGTGTTCTACCTTAATGGAATAATGTTGTTGTTCTCTCTTC 2180
2285 CTCCATCCAGAAATCTTTTGAAGAGCTATCCATGTCATAGTTGTGCTATAGGA 2344
2181 AGCAACTCCAGAAAAACCTTCTCAAGAGGCTATCCATGTCATTAGTTGTGTTATAGGA 2240

QY	2345	CAAGTCTGAATGGGAACCTGAGATTGGTTGGATCTATGATCTGTTCACAGAGATATTTCT	2404
Db	2241	TAAGTCAGATTGGGAATGGAGATTGGATGGATCTATGGTTCTGTGACAGAGATATTTCT	2300
QY	2405	TACTGGATTCAAGATGCACGCAAGAGCTGGGGTTTCAGTCTATTTGATGCCCAAGGCC	2464
Db	2301	GACTGGTTTCAAAATGATGCCCGTGGATGGGATCCATTACTGATGCTTAAGTTCC	2360
QY	2465	AGCTTTCAGGAGATCGGCCCATCATCTTTTCAGATCTGTGAACCAAGTGTGGGTG	2524
Db	2361	AGCTTTCAGGAGTCTGCTCCATCAATCTTTTCAGATCTGTGAACCAAGTGTGGGTG	2420
QY	2525	GGCTCTCGGTTCTGTTGAATTTCTTTTCAGCGGATTCGCCCTATGTTGATGGCTACGG	2584
Db	2421	GGCTTTAGTTTCAGTTGAGATTTCTCTTCAGTCGGCAATTCCTATATGATATGGTTACAA	2480
QY	2585	AGGGCGCTCAAGTTCTCTGGAGAGATTCGCTTTACATCAACACACCATTTACCCACTAC	2644
Db	2481	TGGGAGGCTAAATTTCTTGAGAGGTTTTCGATGTGAACACCATCTACCTATCAC	2540
QY	2645	CTCTCTCCGCTTCTAGTCTATTCTATATGCTGCTATCTGTCTCTCACTGGAAAGTT	2704
Db	2541	CTCCATCTCTCTCATGTATTGTATGATTCAGTTCGAGCCGTTTCTCTTCCACCAACGAT	2600
QY	2705	CATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTTGCGCTCTTCTTTCAT	2764
Db	2601	TATTAATCTCTAGATTAGTAACATTTGCAAGTATATGGTTTCTGTCTCTTCTCTCCAT	2660
QY	2765	TTTCCGCACTGATATCTTTGAGATGAGTGGAGTGGTTGTCATTTGACGATGCTGGAG	2824
Db	2661	TTTCCGCACTGATATCTTTGAGATGAGTGGAGTGGTTGTCATTTGACGATGCTGGAG	2720
QY	2825	GAATGAACAGTTCTGGGTCTATGAGATGATCTCTGCAATCTGTTTGGCGTTTTCAGG	2884
Db	2721	AAACGAGCAGTTTGGGTCTATGAGATGATCTCTGCAATCTGTTTGGCGTTTTCAGG	2780
QY	2885	TCTTCTGAAGTGCTTGGCGGTATGCACCAACTTCTACTGTCACCTCAAGGCTTAATGA	2944
Db	2781	TATCTCTAAGTCTTGGCGGTATGACACAACTTCTACAGTTTCTCAAAAGCTTCAGA	2840
QY	2945	CGAAGAGCGACTTTGCTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCCCTCC	3004
Db	2841	TGAAGACGAGACTTTGCTGAGCTCTACTGTTTCAATGGACAACTTCTGATTCGCC	2900
QY	3005	GACGACCAATTTGATCATTAACATGGTTGGTGTCTGCTGGACCTCTCAGCCCATCAA	3064
Db	2901	AACGACGCTGCTCATTTGTAAACTTAGTGGGAGTTCTTTCAGGAGTCTCTTATGCTATCAA	2960
QY	3065	CAGTGGTTTACCAATCATGGGGCGGCTCTTTGGGAAGCTCTTCTTGGCTCTGGGTGAT	3124
Db	2961	CAGTGGATACCAATCATGGGACCACTCTTTGGTAAAGTTGTTCTTGGCTCTGGGTGAT	3020
QY	3125	TGTTCACTATACCAATCTCAAGGGTCTTATGGGAGGCAAAACCGCACACCGACCAT	3184
Db	3021	TGTTCACTATACCAATCTCAAGGGTCTTATGGGAGGCAAAACCGCACACCGACCAT	3080
QY	3185	TGTTCACTATACCAATCTCAAGGGTCTTATGGGAGGCAAAACCGCACACCGACCAT	3244
Db	3081	TGTTCACTATACCAATCTCAAGGGTCTTATGGGAGGCAAAACCGCACACCGACCAT	3140
QY	3245	TCCATTCACTACCGTCTCGCTGGGCCCAAAATCCAACTGTGCACTCACTGCTAGGA	3304
Db	3141	TCCATTCACTACCGTCTCGCTGGGCCCAAAATCCAACTGTGCACTCACTGCTAGGA	3200
QY	3305	A 3305	
Db	3201	A 3201	

RESULT 13
AF027174
LOCUS

DEFINITION

Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B) mRNA, complete cds.

AF027174.1

GI:2827142

ACCESSION

AF027174

VERSION

AF027174.1

KEYWORDS

Arabidopsis thaliana

SOURCE

Arabidopsis thaliana

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 3682)

Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., Camillieri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J., and Williamson, R.E.

Molecular analysis of cellulose biosynthesis in Arabidopsis

Science 279 (5351), 717-720 (1998)

98111412

9445479

2 (bases 1 to 3682)

Arioli, T.

Direct Submission

Submitted (29-SEP-1997)

Plant Science Centre, Australian National University, Acton, Canberra, ACT 200, Australia

Location/Qualifiers

1..3682

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

1..3682

/gene="Ath-B"

256..3453

/gene="Ath-B"

/note="RSMI-like"

/codon_start=1

/product="cellulose synthase catalytic subunit"

/protein_id="AAC39336.1"

/db_xref="GI:2827143"

/translation="MESSEGTAGPKMKNIVPQTQICSDNVNKTVDGDFVACDICSF

PVCPCYVERKQDNQSPQCKTKLKGSPAIPEGKDEGLDGVGTEVFNYPQKEK

ISERMLGHLTRGKEEMGEQYDKEVSHNHLPLTSPQDTSGESAASPERLSVST

IAGKRLPSSSDVNVQSNRRIVDPVGLGNVAKWERVDGKQKNTGFTVQAASER

CGVDIADTADALNDARQPLPSRRINPFYMWIMLRVILCLFLHYR

ITNPFAFALWLVSVICETWFWALSDLDQPKWFFNRYDLRLALRYDREGPSQ

LAAPDIFVSTVDPLKEPLTANTVLSILAVDYFVKVSCYFDDGAMLSFESLAT

SEFARKVPCKYSIEPRAPETFAAKIDYDKQVTSFVKDRAMKREVEEFKIRI

NALYSKALCKPEGWVMDGTPWPNNTGDHPGMLOVFLGONGGLDARGNELPLRVV

SREKRPQHHKAGAMNALVRVAVLTNGPFIINLDCDHVNNKSKALREACFLMDP

NLGHQVCYVQPPQPDGIDKNDRYANRTVFEDINLGLDSIQGVVYVGTGCVNRTA

LYGPEPLKVKHKKPSLLSKLGGSRKKSKAKKESDKKSGRHTDSTVPVFNLDIE

EGVEGAFDDEKALLMSQSLERKFGQSAVFASVTLMEGVPPSPATPENLKEAIV

ISCYEDKSDMGMEIGWYGVSTEDILTGFMAHGWRSIYCMPLKPAFGKSAATLS

DLNQLRWALGSVEILLPSRHCPIWYNGRLKFLERFAYVNTTTPITSTPLMYCT

LLAVCLTNQFLIPQISNIASIFWLSIFLSIFATGILEMPSGVGIDENWNEQFWI

GVSAHLFANFQGIKVLGADTNFTVTSKASDDGFAELYLFLFWITLLIPPTLLI

VNLGVVAGVSIAINSQVSNGLPFGKLFAPFVWVIVHLVYPLKGLMGRQNRTPITVVV

WSVLLASIFSLMWLRIDPFTSRVTGPDILECGINC"

ORIGIN

Query Match

Best Local Similarity

Matches 2348; Conservative

46.2%; Score 1675.2; DB 8; Length 3682;

71.4%; Pred. No. 0;

Mismatches 868; Indels 72; Gaps 8;

QY 30 TCGGTGGGGTTGATCGCTCTGCCGCGCATGACGCGGACGCGGACGCGCTCAAGTCC 89

Db 229 TCGGTGTTGAAGCACTAAGTGACAAATCGAAGGGGAAACCGCGGGAAGCCG 288

QY 90 GGGAGGCAAGGGCCGGGACGTGTGCCAGATTCGCCCGACGCGCTGGGACCAACGTTG 149

Db 289 ATGAAGAACAATTGTTCCGACAGCTTGCAGATCTGTAGTGACATGTTGCAAGACTGT 348

QY 150 GACGGCAGCTTTCACCGCTCGAGCTGTCCGCTCTCCGCTCTCCGCTCTCCGCTCTAC 209

349 GATGGAGATCGTTTGTGGCTTGTGATATTTGTTCAATCCCAAGTTTGTGGCGCTTGCTAC 408
Db
210 GAGCAGCGGAGAGAGGACCCAGGCTGCTCCAGTGCAGACCAAGTACAGCGC 269
Qy
409 GAGTATGAGAGGAAGATGGAATCAATCTGTCTCAGTGCAGAAACCAAGATACAAGAG 468
Db
270 CACAGAGGAGCCAGCGATCCGCGGAGGAGGAGCGACACTGATGCCGATGATGGT 329
Qy
469 CTCAGAGGTAGTCTCTCTATCTCTGTGATTAAGACGAGATGGCTAGTGTGATGAGGT 528
Db
330 A---GTGACTTCAACTACCTGCTCATCTGGCACTGAGGACCAAGAGCAGAGATGCTGAC 386
Qy
529 ACTGTTGAGTTCACTACCT-----CAGAAGCAGAAATTTTCAGAG 570
Db
387 AGATCGGAGCTGGCGCATGAACACCGGGGAGTGGCATGTTGGCCACCCCAAGTAT 446
Qy
571 CGGATGCTGGTTGGCATCTTACTCTGTGGGAAGGAGAGAAATGGGGGAACCCCAAGTAT 630
Db
447 GACAGTGGGAGATCGGCTCTCCAAATATGACAGTGGAGAGATCCCTAGGGGATACGTC 506
Qy
631 GATAAAGAG-----TCTCTCAATCATCTTCTCTCTCAGGACGACAAAG-- 679
Db
507 CTTTCACTCACCACAGCCAGATGTTCAGGAGAAATCCCTGGAGCTTGGCTGATCATCAC 566
Qy
680 -----ATACTTCAGGAGATTTTCTGCTCACCTGAAAGCGCTC 720
Db
567 ATGATGTCCCTACGGGGAACATCAGCAGACGTGCTCCGTTTCCT---ATGTGATCAT 623
Qy
721 TCTGTATCTTCTACTATCGCTGGGGAAAGCGCTTCCCTATTATCATAGATGCAATCAA 780
Db
624 TCACCAATCCGTCAAGGAGTGTCTCGGCAGTATTGGGAATGTTGCTCGAAAGAGAGA 683
Qy
781 TCACCAATFAGAGGATTTGGATCTGTTGGACTCGGGAATGAGCTTGAAGGAGAGA 840
Db
684 GTTGATGGCTGGAATTAAGACAGGACAAGGGTGGATCCCATGATTAATGGGACAAGC 743
Qy
841 GTTGATGGCTGGAATTAAGACAGGACAAGAAATACTGCTC-----TGTGAGCAGC 891
Db
744 ATTGCTCCCTCTGAAGTGGGAGCTACTGACATCATCATCTACTGATCAACATG 803
Qy
892 CAGGCTCTTCTGAAGAGGTGGAG---TAGATATTGATGCCAGCAGATATCCTAGCA 948
Db
804 GAGACGCTTTACTGATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCAATGCT 863
Qy
949 GATGAGCTCTGTAATGATGAGAGCAGGAGCAGCTCTGTCAAGAAAGTTTCAATTCCT 1008
Db
864 TCCTCCAAATAAATCCCTACAGAAATGGTCAATGTTCTGGGGTGGTGTCTTAAGCATC 923
Qy
1009 TCATCCGATCAATCTCTACAGAAATGGTTATTATGCTGGCGCTTGTATCTTGTCTC 1068
Db
924 TTCCTGCACTACGCTCTCAAAATCCCTGCGTAAATGCATACCCACTGTGGCTTTTATCT 983
Qy
1069 TTCTTGATTTACCGTATAACAAACCCAGTGCAGAAATGCTTGTCTATGGCTGGTCTCT 1128
Db
984 GTTATATGTGAGATTGGTTTGTCTTTATCTCTGGATCTGATCAGTTCCTCCAGTGGTTT 1043
Qy
1129 GTGATATGTGAGATCTGGTTGCTCTATCTCTGATTTTGGATCAGTTCCTCCAGTGGTTT 1188
Db
1044 CCAATCAACCGGGAGACTCTCTGATAGATCGGCTTTAAGTATGACCGAGAGAGTGAA 1103
Qy
1189 CTTGTGAACCGTGAACCTACTCGACAGGCTTCTTTAAGATATGATCGTGAAGGTGAG 1248
Db
1104 CCGTCTCAGTGGCTGCTGTTGACATATTTCTGATCAGTCCAGCCCTTCAAGAGGCA 1163
Qy
1249 CCATCAGATTAGTCTGTTGACATTTTCTGAGTACTGTTGACCCCTTGAAGAGGCA 1308
Db
1164 CCTATCGTCACTGCGCAACACTGTGCTATCCATTTCTGCTGTGATTAATCCGTCAGCAAG 1223
Qy
1309 CCCCTTGTGACAGCCCAACAGTGTCTCTATCTCTGCTGTTGACTACCCAGTTGACAA 1368
Db
1224 GTCTCTGCTATGATCTGATGACGAGCTTCAATGCTGACTTTGACGATTTGGCTGAG 1283
Qy
1369 GTGTCCTGTTATGTTTGTGATGATGGTGTCTGCTATGTTATCATTTGAATCATCTTGCAGAA 1428
Db

1284 ACTTCAGAGTTTCTAGGAAATGGGTACCAATTTGTGAAGAGTATGACATTTGAACCAGA 1343
Qy
1429 ACATCAGAGTTTCTCGTAAATGGGTACCAATTTTGCAGAAATATATAGCATAGAGCTCGT 1488
Db
1344 GTCCCGAGTTTACTTTTGGCAGAAATTTGATTACTCTGAAAGACAAAGTCCAGCTTCA 1403
Qy
1489 GCACCAGAATGCTACTTTGCTCGAAATAGATTAATCTTGAAGGATAAAGTTTCAGACATCA 1548
Db
1404 TTTGTTAAAGACCGCGGCTCATGAGACAGATATGAAGAAATTTAAATTCAGATAAAT 1463
Qy
1549 TTTGTCAAAGATCGTAGAGCTATGAAGAGGGAATATGAGAAATTTAAATTCGAATCAAT 1608
Db
1464 GCCTAGTTTCTAAAGCATTTGAAAGTCCCGAGGAAAGGATGGAATCATGCAAGATGGCACA 1523
Qy
1609 GCATTTGTTTCCAAAGCCCTAAATGCTCTGAAGAGGGTGGGTATGCAAGATGGCACA 1668
Db
1524 CAAATGCGGAGGAAACAAATACAGGGAATCATCTCGGAATGATTCAGGTTTTCTTGTCTCAC 1583
Qy
1669 CCGTGGCTGGAATTAATACAGGGGACCATCCAGGAATGATCCAGGTCTTCTTAGGCGCA 1728
Db
1584 AGTGGTGGCTTGATACTCAGGGTAAATGAGCTCCCGCTTTTGTATTTATGCTCTCGTAA 1643
Qy
1729 AATGGTGGACTTGTATGAGAGGGAATGAGCTCCCGGTTTGGTATATGTTTCTCGAGAA 1788
Db
1644 AAGCGTCTCTGGGTTCAGCACCACAAAGAGGCTGGTCCCATGAATGCCCTTGTGCTGTC 1703
Qy
1789 AAGCGACACAGGATTCAGCACCACAAAGAGGCTGGTCTATGAATGCTGCTGAGAGTT 1848
Db
1704 TCAGCTGTCTTACTTAATGGAACAATACATGTTGAATCTTGAATTTGATCTACTACATCAAC 1763
Qy
1849 TCAGCAGTTCTTACCAATGGACCTTTTCACTTGAATCTTGAATTTGATCAATTACATTAAT 1908
Db
1764 AACAGAAAGCTCTCGAGAAAGTATGCTCTCTAATGATCCAAACCTTAGTCTCGCAGAA 1823
Qy
1909 AACAGAAAGCTTTAAGAGAGCAATGTGCTTCTGATGACCCCAACCTCGGGAAGCAA 1968
Db
1824 GTCTGTTATGTGAGTTTCCCAAAAGGTTTGAATGGGAATGATAGGAATGATGATGCA 1883
Qy
1969 GTTGTGTTATGTTTCAGTTCCCAAAAGATTTGATGGTATCGATAAGAACGATAGATATGT 2028
Db
1884 AACAGAAAGCTGCTTTTGTATTAATCTTGAAGGGCTTGAAGGCTTTCAGGAGCA 1943
Qy
2029 AATCGTAATACCGTGTCTTTGATATTAATCTGAGAGGTTTAGATGGATTCAGAGACT 2088
Db
1944 GTTATGTGGAACTGCTGTTTTCACAGAAACAGCTATCTATGTTGTTATGAGCCCCCA 2003
Qy
2089 GTATATGTCGAACTGATGTTTTCACAGAAACAGCTATTAACGCTTATGAACTTCA 2148
Db
2004 ATTAAGG-----CGAAGAACCCAGGTTTCTTGGCATCACTATGTTGGGGGCAAGAAAG 2057
Qy
2149 ATAAAGTAAACACAAAGAACCAAGTCTTTTATCTAAGCTCTGTGGTGGATCAAGAAAG 2208
Db
2058 GCAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAGTGCAGAACAGCATGTGACAGT 2117
Qy
2209 AAGAAATTCAGAAAGCTAAGAAAGAGTCCGACAAAGAAATACAGCAGCATACTGACTCA 2268
Db
2118 TCTGTTCCAGTATTAATCTCGAAGACATAGAGGAGGTTGTTGAAGGTGCTGGGTTGAT 2177
Qy
2269 ACTGTTCTGTATTAACCTCGATGACATAGAGAGGAGTTGAAGGTGCTGGTGGTGGT 2328
Db
2178 GATGAGAAATCAGTCTCATGTCTCAAAATGAGCTTGAAGAGAGATTTGGCAGTCAAG 2237
Qy
2329 GATGAAAGGCGCTCTTAATGTGCAAAATGAGCCTGGAGAGCGATTTGGAACGTCTGCT 2388
Db
2238 GCATTTGCTGCTCCTCTGATGGAATATGTTGTTGTTCTCAGTCTCTCAGTCTCCAGTCA 2297
Qy
2389 GTTTTGTGTTCTTACCTTAATGAAATGCTGGTGTCTCTCTCAGCACTCCAGAA 2448
Db
2298 TCTCTTTGAAAGAGCTATCCATGTCATAGTTTGTGCTATGAGGACAAAGTCTGAATGG 2357
Qy
2449 AACCTTCTCAAGAGGCTATCCATGTTCATTTAGTTTGTGTTATGAGGATAAGTCAAGTTGG 2508
Db

QY 2358 GGAACCTGAGATTGGTTGGATCTATGATCTGTACAGAGATATCTTACTGTGATCAAG 2417
Db |||||
QY 2509 GGAATGGAGATTGGATGGATCTATGTTCTGTGACAGAGATATCTGACTGGTTCAG 2568
Db |||||
QY 2418 ATGACGCAAGAGCTGGCTGCTAGTCTATGATGCCAAGCCGACGCTTTCAGAGGA 2477
Db |||||
QY 2569 ATGCATGCCGCTGGATGGGATCCATTACTGATGCCCTAAGCTTCAGAGCTTCAAGGCT 2628
QY 2478 TCTGCCCCCACTCAATCTTTTCAGATCTCTGACCAAGTCTCGGTGGGCTCTCGTTCT 2537
Db |||||
QY 2629 TCTGCTCTCATCAATCTTTTCAGATCTCTGACCAAGTCTGAGTGGGCTTTAGGTTCA 2688
QY 2538 GTTGAATCTTTTCAGCCGCAATCCCTCTATGATGATGGCTACGGAGGCGCCTCAAG 2597
Db |||||
QY 2689 GTTGAATCTCTTCAGTCCGATCTCTATATGATGATGTTTACAAATGGAGGCTAATA 2748
QY 2598 TTCTCTGGAGATTTGCTTACATCAACACCAATTTACCCATAACCTCTCTCCGCTT 2657
Db |||||
QY 2749 TTCTCTGGAGATTTGCTTACATCAACACCAATTTACCCATAACCTCTCTCCGCTT 2808
QY 2658 CTAGTCTATTGATATTTGCTCTGATCTGCTGCTCACTGGAAGTTTCATGCGCAG 2717
Db |||||
QY 2809 CTCATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2868
QY 2718 ATTAGCACTTGGCCAGTATCTGGTTTCAATGGGCTCTTCTTTTCAATTTTGGCCACTG 2777
Db |||||
QY 2869 ATTAGTAACTTGAAGTATATGTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2928
QY 2778 ATCTTGAATGAGTGGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2837
Db |||||
QY 2929 ATACTAGAAATGAGTGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2988
QY 2838 TGGGTCATTTGAGGATCTCTGACATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2897
Db |||||
QY 2989 TGGGTCATTTGAGGATCTCTGACATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3048
QY 2898 CTGCGGCTATCGACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2957
Db |||||
QY 3049 CTGCGGCTATTTGACACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3108
QY 2958 TTTGCTGAGCTCTGATGTTTGAAGTGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3017
Db |||||
QY 3109 TTTGCTGAGCTCTGATGTTTGAAGTGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3168
QY 3018 ATCATTTACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3077
Db |||||
QY 3169 ATTGTAACTTTAGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3228
QY 3078 TCATGGGCGCCCTCTTTGGGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3137
Db |||||
QY 3229 TCATGGGCGCCCTCTTTGGGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3288
QY 3138 CAATTTCTCAAGGCTCTTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3197
Db |||||
QY 3289 CTTTCTCTCAAGGCTCTTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3348
QY 3198 GCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3257
Db |||||
QY 3349 TCTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3408
QY 3258 GCTCTGCTGCGCCCAATATCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3305
Db |||||
QY 3409 CGAGTCACTGCGCCGACATCTGGAATCTGGAATCACTGTTTGAA 3456
Db |||||

RESULT 14
AR267559
LOCUS 3614 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 9 from patent US 6495740.
ACCESSION AR267559
VERSION AR267559.1 GI:29697635
KEYWORDS
SOURCE Unknown.

ORGANISM

Unknown.

REFERENCE

Unclassified.

1 (bases 1 to 3614)

AUTHORS Arioli, A., Williamson, R. E., Betzner, A. S. and Peng, L.

TITLE Manipulation of cellulose and/or beta.-1,4-Glucan

JOURNAL Patent: US 6495740-A 9 17-DEC-2002;

FEATURES Location/Qualifiers

source

1..3614

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 46.2%; Score 1673.6; DB 6; Length 3614;

Best Local Similarity 71.4%; Pred. No. 0;

Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;

QY 30 TCGGTGCGGCTTGGATCGCTCTGCGCGGCATAGTGAACAAATGGAATCGAAGGAGAAACCGGGAAGCCG 89

Db 190 TCGGTGTTGGAAGCAACTAAGTGACAAATGGAATCGAAGGAGAAACCGGGAAGCCG 249

QY 90 GGGAGCAGCGGCGCGGAGACGTGTGCGAGATCTGCGCGCAGCGGCTGGGACACAGTTG 149

Db 250 ATGAAGAACATTTGTTCCGCGAGACTTGCAGACTCTGTAGTGAACAATGTTGGCAAGACTGTT 309

QY 150 GACGCGAGCTCTTCAACGCGCTGCGAGCTCTGCGGCTTCCGCTCTGCGGCTTCCGCGCTGCTAC 209

Db 310 GATGAGATTCGTTTGTGGCTTGTGATTTGTTTATTTCCAGTTTGTGCGGCTTGTCTAC 369

QY 210 GAGCAGCAGCGCAAGGAGGCGCACCCAGGCTTCCCTTCCAGTGAAGACCAAGTACAAGCGC 269

Db 370 GAGTATGAGAGAAAGATGGGAATCAATCTTGTCTCAGTGAACCAACAGATACAAGAGG 429

QY 270 CACAGGAGGCGCCAGCGATCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 329

Db 430 CTCAAAGGTAGTCTGCTGCTTATCTCTGCTGATAAAGACGAGGATGGCTTAGCTGATGAAGT 489

QY 330 A---GTGACTTCAACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 386

Db 490 ACTGTTGAGTTCAACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531

QY 387 AGGATCGCAGCTGCGCATGAACACCGGCGGCGAGTGGCAATGTTGGCCACCCCAAGTAT 446

Db 532 CGGATGCTTGTGCTGCTTACTCTGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 591

QY 447 GACATGCGAGATCGGCTCTTCAAGTATGACATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 506

Db 592 GATAAAGAGG-----TCTCTCAATCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640

QY 507 CTTTCACTCACCACGCGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 566

Db 641 -----ATCTTTCAGGAGGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681

QY 567 ATGATGCTCCCTTACCGGGAACATCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623

Db 682 TCTGATCTTCTTACTACTCTGCTGCGGGAAGCGCTTCTCTTATCATCAGATGCTCAATCA 741

QY 624 TCACCAATTCGCTCAAGGAGGAGTCTTCCGCGAGTATTTGGGAATTTGCTTGGGAGGAGGAGGAGG 683

Db 742 TCACCAATATAGAGGAGTGTGGATCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801

QY 684 GTTGATGCTGGAATATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 743

Db 802 GTTGATGCTGGAATATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 852

QY 744 ATTGCTCCCTTACCGGGAACATCAGCAGAGCTTACTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 803

Db 853 CAGGCTGCTTCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 909

QY 804 GAAGAGCTTTTACTGATGATGAATCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 863

Db 910 GATGAGGCTCTGCTGATGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 969

QY 854 TCCTCCAAATAAATCCCTCCTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923

Db 970 TCAACGGATCAATCTTACAGAAATGGTTATATGCTGGGCTTGTATCTCTTGTCTC 1029
QY 924 TTCCTGACCTACCGTCTCAAAATCCCTGTCGTAATGCAATACCACTGTCGCTTTTATCT 983
Db 1030 TTCTTGCAATACCGTATACAAACCCAGTCGCAAAATGCCCTTTGCTCTATGGCTGGTCTCT 1089
QY 984 GTTATATGAGATTTGGTTTGTCTTATCTGATGACTGATCAGTTCCCGAGTGGTTT 1043
Db 1090 GTGATATGTGAGATCTGGTTTGGCTTATCTGGAATTTGGATCAGTTTCCCAAGTGGTTT 1149
QY 1044 CCAATCAACCGGAGACCTACCTTTGATAGACTGGCTTTTAAAGTATGACCCGAGAGGTGAA 1103
Db 1150 CCGTGAACCGTGAACCTTACCTCGACAGCTTTCCTTTAAGATATGATCGTGAAGGTGAG 1209
QY 1104 CCGTCTCAGTTGGCTGCTGTGTGACATATTTGTGATGACAGTCGACGCCCTTGAAGGACCA 1163
Db 1210 CCATCAAGTTAGCTGTGCTGTGACATTTTGGTGAATCTTTCACCCCTTGAAGGACCA 1269
QY 1164 CCTATCGTCACTGCCAACACTGTGCTATCCATTTCTGCTGTGATATCCGCTGGACAAG 1223
Db 1270 CCCCCTGTGACGCAACACAGTGTCTCTATTTCTGGCTGTGATACCCAGTTGACAG 1329
QY 1224 GTCTTGTGATGATCTGATGACGAGCTTCAATGCTGACTTTTGGACGATTTGGCTGAG 1283
Db 1330 GTGCTCTGTTATGTTTCTGATGATGCTGCTGCTATGTTATCAATTTGAATCACITGCGAGAA 1389
QY 1284 ACTTCAGAGTTTGTAGAAATGGGTACCAATTTGTGGAAGAGTATGACATTTGAACCCAGA 1343
Db 1390 ACATCAGAGTTTGTCTGTAATGGTACCAATTTTGCAGAAATATAGCATAGAGCTTCGT 1449
QY 1344 GCTCCCGAGTTTACTTTTCCAGAAATTTGATTTACCTGGAAGACAAAGTCCAGCCTTCA 1403
Db 1450 GCACCAAGATGGTACTTTGCTGCGAAATATAGATTTACTTGAAGGATAAAGTTTCAGACATCA 1509
QY 1404 TTTGTTAAAGACCCCGGGCCATGAAGAGAAATATGAAGAAATTTAAATCAGGATAAAT 1463
Db 1510 TTTGTCAAAGATCGTAGAGCTATGAAGGGAATATGAGGNAATTTAAATCCGAATCAAT 1569
QY 1464 GCCCTAGTTTCAAGGCAATGAAGTCCCGAGAGAGATGGATCATGCAAGATGGCACA 1523
Db 1570 GCACTTGTTCCTAAAGCCCTAAAATGCTCTGAAGAGGGTGGGTATGCAAGATGGCACA 1629
QY 1524 CCATGGCCAGAAACATACAGGATCATCTCGAATGATCAGGTTTTCCTTGTGTCAC 1583
Db 1630 CCGTGGCTTGAATAATACAGGGGACCATCCAGGAATGATCCAGGTCCTTTAGGGCAA 1689
QY 1584 AGTGGTGGCTTGTATCTGAGGGTAATGAGCTCCCGGCTTTTGTATGTTGTCTCGTGA 1643
Db 1690 AATGTTGACTTGTATGACAGAGGCAATGAGCTCCCGGCTTTGGTATATGTTTCTCGAGAA 1749
QY 1644 AAGCGTCTGGGTTCCAGACCAACAGAGGCTGGTGCATGAATGCCCTTGTGCTGTC 1703
Db 1750 AAGCGACCGAGTTTCCAGCACCAAAAGGCTGGTGTATGAATGCACCTGGTGAGAGTT 1809
QY 1704 TCAGCTCTCTTACTTAATGACAAATACATGTTGAAATCTTGATTTGATCACTACATCAAC 1763
Db 1810 TCAGAGTTCTTACCAATGGAACCTTTCAITTTGAAATCTTGATTTGATCATATCAAT 1869
QY 1764 AACAGCAAGGCTGTCGAGAGAGCTATGTGCTTCTTAATGGAATCCAAACCTAGGTCCGCAA 1823
Db 1870 AACAGCAAGGCTTAAAGAGAGCAATGTGCTTCTGTGAGACCCAAACCTCGGGAAGCAA 1929
QY 1824 GTCTGTTATGTCAGTTCCTCAACAAAGGTTTGAATGGATTTGATAGGAATGATCGATATGCA 1883
Db 1930 GTTGTGTTATGTCAGTTCCTCAACAAAGATTTGATGATTCGATATGAAGAGATAGATGCT 1989
QY 1884 AACAGGAACACTGTCTTTTGTATTAATTAACCTTGAAGGCTTGAAGGCAATCAAGGACCA 1943
Db 1990 AATCGTAATACCGTGTCTTTGATATTAACCTTGAAGGTTTAGATGGGATTCAGGACCT 2049
QY 1944 GTTATATGGAACCTGGTTGTGTTTCAACAGACAGCTATCTATGTTGTTATGAGCCCCCA 2003

2050 GTATATGTCGAACTGGATGTGTTTTCACAGAAACAGCAATATACGGTTATGAACCTCCA 2109
2004 ATTAAAGG-----CGAAGAGCCAGGTTTCTTGGCATCACTATGTCGGGCAAGAAAG 2057
2110 ATAAAGTAAACACAGAAAGCCAGTCTTTTATCTAAGCTCTGTGTGATCAAGAAAG 2169
2058 GCAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAGTCGAACAAAGCATGTGGACAGT 2117
2170 AAGAAATCCAAAGCTAAGAAAGAGTCCGACAAAGAAATCAGCGAGGCATCTGACTCA 2229
2118 TCTGTTCCAGTATTTCAATCTCGAAGACATAGAGAGGGTGTTCGAAGTCTGGTGTGAT 2177
2230 ACTGTTCTCTGATTTCAACCTCGATGACATAGAAAGAGGGAGTTGAAGGTGCTGGTTGAT 2289
2178 GATGAGAAATCAGTCTCTCATGCTCAAAATGAGCTTAGAGAAAGAGATTTGGCCAGTCAGCA 2237
2290 GATGAAGAGCGCTCTTAATGTCGAATATGAGCTCGAAGAGGATTTGACAGTCTGCT 2349
2238 GCATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTGTTCCTCAGTCTCTCCATCCAGAA 2297
2350 GTTTTTGTGCTTCTACCTTAATGGAATATGGTGGTGTTCCTCCTTCAGCAACTCCAGAA 2409
2298 TCTCTTTGAAAGAGCTATCCATGTCATTAAGTTGTGGCTATGAGGACAACTCTGAATGG 2357
2410 AACTTTCTCAAAGAGGCTATCCATGTCAATAGTTGTGGTTATGAGGATAAGTCAGATTGG 2469
2358 GGAACCTGAGATTTGTTGGATCTATGATCTGTCAAGAAATATCTTACTGATTTCAAG 2417
2470 GGAATGAGATTTGATGATCTATGTTCTGTGACAGAAATATCTGACTGGTTCAAA 2529
2418 ATGCAGCCAAAGAGCTGGCGTTTCAGTCTAATTTGATGCCAAAGCCAGCTTTCAAGGGA 2477
2530 ATGCATGCCGCTGGATGGCATCCATTTACTGATGCTTAAGCTTTCAGCTTTCAAGGTT 2589
2478 TCTGCCCCCATCAATCTTTGAGATCTGTGAAACCAAGTCTGCGTGGCTCTCGTCT 2537
2590 TCTGCTCTATCAATCTTTTCAGATCTGTGAAACCAAGTCTGAGTGGCTTTAGGTTTCA 2649
2538 GTTGAATCTTTTTCAGCGGCAATGGCCCTTATGGTATGGCTACGAGAGGCGCTCAAG 2597
2650 GTTGAGATTTCTTTTCAGTCGGCATTTGCTCTATATGTTATGTTTACATGGAGGCTAAAA 2709
2598 TTCCTGAGAGATTCGCTTACATCAACACCACTTTACCCACTTAACCTCTCTCCGCTT 2657
2710 TTTCTTGAGAGGTTTGGCTATGTGAAACCAACCATCTACCTATCACTCCATTCCTCTT 2769
2658 CTAGTCTATTTGATATTTGCTGCTATCTGTCTGCTACCTGGAAGTTTCACTATCCAGAG 2717
2770 CTCATGTTATTTGATCTAGCGGTTTGTCTCTTCCAAACCCAGTTTATTTATTCCTCAG 2829
2718 ATTAGCAACTTGGCCAGTATCTGTTTCAATTTGCGCTCTTCTCTTCAATTTTTCGCCACTGGT 2777
2830 ATTATGAACATTTGCAAGTATATGTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
2778 ATCTTTCAGATGAGGTTGGTGTGTTGCAATTTGACAGAGTGGTGGAGGAATGAACAGTTTC 2837
2890 ATACTGAATGAGTGGAGTGGCTAGGCATAGACGAATGGTGGAGAAACAGCAGATTT 2949
2838 TGGTCAATGGAGTATCTCTGACATCTGTTTGGCGCTCTTTCAGGCTCTTCTCAAGGTG 2897
2950 TGGGTCAATGGTGGAGTATCGGCTCAATTTATGCTGTGTTTTCAGGTTATCTCTCAAGATC 3009
2898 CTTGCCGTTATCGACACCACTTCACTGTCACTCAAGGCTAATGACAGAGAGGCGAC 2957
3010 CTTGCCGTTATGACAAACTTCAAGTTTACCTCAAAAGCTTCAGATGAAGACGGAGAC 3069
2958 TTTGCTGAGCTCTACATGTTTCAAGTGAAGAGCGTTCTCATCTCCCTCCGACGACATTTTG 3017
3070 TTTGCTGAGCTCTACTTGTTCAAATGGACAACTTCTGATTTCCGCGCAACAGCCTGCTC 3129
3018 ATCAATTAACATGGTGTGTTGCTGGCACCTCTACGCTACCAAGTGGTTTACCAA 3077
3130 ATTGTAAACTTAGTGGAGTTTGTGAGGAGTCTCTTATGCTATCAACAGTGGATACCA 3189

QY 3078 TCATGGGGCCGCTCTTTGGAGCTCTTCTTGGCTTCTGGTGAATGTTCACTATAC 3137
 Db |||||
 QY 3190 TCATGGGGACCACTCTTTGGTAAATGTTCTTTGGCTTCTGGGTGAATGTTCACTGTATAC 3249
 Db |||||
 QY 3138 CCATTCTCTCAAGGGTCTTATGGGCGAGCAAAACCGCACACCGACGATTGTCATCTGCTGG 3197
 Db |||||
 QY 3250 CCTTCTCTCAAGGGTCTTATGGGTCGACAGAACCGGACTCTCTACCATTTGTTGTTCTGG 3309
 Db |||||
 QY 3198 GCTGCTCTCTGCTTCTATCTTCTCTGCTGCTGGGTTGCTGTTGATCCATTCACTACC 3257
 Db |||||
 QY 3310 TCTGTTCTCTTGGCTTCTATCTTCTGTTGTTGTTGTTGTTAGGATTGATCCCTTCACTAGC 3369
 Db |||||
 QY 3258 CTTCTGCTGGCCCAAAATCAAACTCTGTCATCACTAGTACGAA 3305
 Db |||||
 QY 3370 CGAGTCACTGGCCCGGACATCTCGGAATGTGAATCAACTGTGTGAA 3417
 Db |||||

RESULT 15
 AX030946 3614 bp DNA linear PAT 24-NOV-2000
 LOCUS
 DEFINITION Sequence 9 from Patent WO9800549.
 ACCESSION AX030946 AX030960
 VERSION AX030946.1 GI:10278349
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS
 TITLE Manipulation of cellulose and/or beta -1,4-glucan
 JOURNAL Patent: WO 9800549-A 9 08-JAN-1998;
 WILLIAMSON RICHARD EDWARD (AU); PENG LIANGCAI (AU); ARIOLI ANTONIO (AU); UNIV AUSTRALIAN (AU); BETZNER ANDREAS STEFAN (AU); COMWU SCIENT IND RES ORG (AU)
 COMMENT On Oct 15, 2002 this sequence version replaced gi:10278363.
 FEATURES
 Location/Qualifiers
 1..3614
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /strain="COLUMBIA"
 /db_xref="taxon:3702"
 /clone="ATH-B"
 217..3414
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC09627.1"
 /db_xref="GI:10278350"
 /db_xref="REMBL:CAC09627"
 /translation="MESGETAGKPMKNIVPTQICSDNVGKTVGDRFVACDICSF
 PVRPCYERKDNQSCPCQKRYKRLKSPALPGDKDEGLADEGTVFENYPOKEK
 ISEMLGHLTRGKEEMGEPEQYDKVEVSHNPLRLTSQDTSGFSAASPERLSVST
 IAGGRPLPSDVMNRRIVDPVGLGNVAKWVKVQKQKNTGPTVQAAER
 GVDIDASTDILADEALNDAROLLSKVSIPSRINPBYMLRLVILCLFLHYR
 ITNPVPNAPALWVSVCEIFALSWILDQPKFPFVNRETYLORLAURYDREGPSQ
 LAVDIFVSTVDPLKEPPLVTANTVLSILAVDPVDKVSXVSDDDGAAMLFSFSLAET
 SEFARKVFPCKKYKSIIEPAPETWPAKIDYLDKQVQTSFVKDRMRKREYEEFKIRI
 NALVSKALKCEEWGMQDGPWPQNTGDHPFLQVGLQNGGLDAEENLRLVYV
 SREKRPQPHKKGAMNALRVSAVLNTGPFILNDCDHYNNKSLAREAMCLMDP
 NLGQVCVQPPQPDGDKNDRYANRNTVFEDINLRGLDGIQGPVYGTGCVENRTA
 LYGEPPPIKVHKPFLSKLSCGSRKKNKAKESDKKSGRHRDTSVPVFNLDIE
 EGVEGAFDDDEKALLMSQMSLEKRFQSAFVATLMEGVPSPATPENFLKAIHY
 ISCGYEDSKDMGEIWTGYGVTEDILTFKMHARGWRSIYCMPLKPAFKGAPNLNS
 DRLNQRLWALSGVEILFSRHCPIWYGNGLKFLERPAYVNTIYPTISPLMYCT
 LLAVCLFTNQFIIPQISNIASFWSLFLSIPTATGILEWRWSGVGIDEWNEOPFWI
 GGVSAHLFAVQGLKVLADITNFTVTSKASDEGDEFAELYLEKWTLLIIPPTLLI
 VNLGVAVAGVAYNSQSWGPLFGKLPFAFWIUVHLYPFLKGLMQRNRTPIVVV
 WSVLIASIFSLWVRIDPFTSRVTPGDPILECGINC"

Query Match 46.2%; Score 1673.6; DB 6; Length 3614;

ORIGIN

Best Local Similarity 71.4%; Pred. No. 0;
 Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;

QY 30 TCGTGCCTGGTGGATCGCTCTGCCGCGCCATCGACGCGACGCGCCCTGAAGTCC 89
 Db |||||
 QY 190 TCGGTGTTGGAAGCACTAAAGTGACAAATGGAATCCGAAGGAGAAACCGCGGAAGCG 249
 Db |||||
 QY 90 GGGAGGACGGGGCGGGGACGTGTGCCAGATCTCGGCGGACGGCTGGGACCACTGTTG 149
 Db |||||
 QY 250 ATGAAGAACATTTGTTCCGACACTTGCAGACTCTGTAGTGACAAATGTTGGCAAGACTGTT 309
 Db |||||
 QY 150 GACGCGGAGCTCTTCAACGCTGCGACGCTCTGCCGCTTCCCGTCTGCGGCCCTGCTAC 209
 Db |||||
 QY 310 GATGAGATCGTTTGTGCTTGTGATATTTGTTCAITCCCAGTTTGTGCGCTTGTCTAC 369
 Db |||||
 QY 210 GAGCAGGCGCAAGGAGGCGCACCCAGGCTCTCCAGTCAAGACCAAGTACAAAGCGC 269
 Db |||||
 QY 370 GAGTATGAGAGGAAAGATGGGAATCAATCTTGTCTCAGTGCACAAACACAGATACAAGAG 429
 Db |||||
 QY 270 CACAGAGGAGCCAGCGATCCGCGGGAGGAGGAGGACGACACTGATCCGATGATGTT 329
 Db |||||
 QY 430 CTCAAAGGTAGTCTCTGCTATTCTCTGCTGATAAAGACGAGGATGGCTTAGCTGATGAAGT 489
 Db |||||
 QY 330 A---GTGACTTCAACTACCTGCTCATCTGGCACTGAGGACCAAGACGAGAAATTCGTAC 386
 Db |||||
 QY 490 ACTGTTGATTCACACTACCT-----CAGAAGGAGAAATTTACAGAG 531
 Db |||||
 QY 387 AGGATGCGCAGCTGGCGATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTAT 446
 Db |||||
 QY 532 CGGATGCTTGGTGGCATCTTACTCGTGGGAGGAGAGGAAATGGGGGAAACCCAGTAT 591
 Db |||||
 QY 447 GACAGTGGCGAGATCGGCTCTCAAAGTATGACAGTGGAGAGATCCCTAGGGGATACGTC 506
 Db |||||
 QY 592 GATAAAGAGG-----TCTCTCACAATCATCTTCTCTGCTCAGGACGAGACAAG-- 640
 Db |||||
 QY 507 CTTTCAGTCACCAACAGCCAGATGTCAGAGAGAAATCCCTGGAGCTTCGCTGATCATCAC 566
 Db |||||
 QY 641 -----ATACCTCAGAGAGTTCCTGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCT 681
 Db |||||
 QY 567 ATGATGCTCCCTACGGGAAACATCAGCAGAGCTGCTCGCTTTTCCCT---ATGTGAATCAT 623
 Db |||||
 QY 682 TCTGTATCTTCTACTATCGCTGGGGGAAAGCGCTTCCCTATTATCATCAGATGCTCAATCAA 741
 Db |||||
 QY 624 TCACCAATTCGTCAGGAGTTCCTCGGCGAGTATTTGGGATGTTGCTCGGAGAGAGAGA 683
 Db |||||
 QY 742 TCACCAATTAGAAGGATTTGTGATCCTCTGAGCTCGGGAATGTAGTTGGAAGGAGAGA 801
 Db |||||
 QY 684 GTTGATGCTGGAAATGAAGCAGGACAAAGGTGCGATTCCCATGACTAATGGGCAAGC 743
 Db |||||
 QY 802 GTTGATGCTGGAAATGAAGCAGGAGAAAGTACTGGTCC-----TGTACAGCAG 852
 Db |||||
 QY 744 ATTGCTCCTCTGAAGGTGCGGCGAGCTACTGACATCGATGCTACTTACTGAAATACAAACATG 803
 Db |||||
 QY 853 CAGGCTGCTTCTGAAAGAGGTGGAG---TAGATATTGATGCCAGCACAGATATCCTAGCA 909
 Db |||||
 QY 804 GAAGACGCTTTACTGAATGATGAAATCTGCCAGCCTCTATCTAGAAAAGTCCCATTCCT 863
 Db |||||
 QY 910 GATGAGGCTCTGCTGAATGACGAAGCGGAGGAGCTTCTGTCAAGGAAAGTTCATATTCCT 969
 Db |||||
 QY 864 TCCTCCAAAATAAATCCCTCAGAGATGCTCATTTCTGCGGTTGGTGTGTTCTAAGCATC 923
 Db |||||
 QY 970 TCATCAGGATCAATTCCTTACAGAAATGTTATATGCTGCGGCTTGTATCTTTGTCTC 1029
 Db |||||
 QY 924 TTCCTGCACTACCGTCTCACAAATCTTGTGCGTAAATGATACACCACTGTGCTTTTATCT 983
 Db |||||
 QY 1030 TTTCTTGCAATTACCGTATATAACAAACCCAGTGCCTTGTCTTCTATGCTGCTGCTCT 1089
 Db |||||
 QY 984 GTTATATGTAGATTTGGTTTGTCTTATCTCTGGATACCTGGATCAGTTCGCCAGTGGTTT 1043
 Db |||||
 QY 1090 GTGATATGTAGATCTGGTTTGGCTTATCTCTGGATTTTGGATCAGTTTCCCAAGTGGTTT 1149
 Db |||||
 QY 1044 CCAATCAACCGGAGACCTACCTTTGATAGATCGCTTTAAGGTATGACCGAGAGGTGAA 1103
 Db |||||

Db	1150	CCGTGTAACCGTGAACCTACCTCGACAGGCTTGTCTTAAAGATATGATCGTGAAGGTGAG	1209
Qy	1104	CCGTCTCAGTTGCTGCTGTTGACATATTTGTTCAGTACAGTCGACCCCTGAGGAGCCCA	1163
Db	1210	CCATCAAGTACGTCTGTGACATTTCTGTCGAGTACTGTTGACCCCTTGAAGGAGCCCA	1269
Qy	1164	CCATCGTCTACTCGCAACACTGTCGTATCCATTTCTGCTGTTGATTTATCCCGTGGACAAG	1223
Db	1270	CCCTTGTGACAGCCACACAGTGTCTTAITCTGGCTGTTCACCTACCAGCTTGACAAG	1329
Qy	1224	GTCTCTGTCTATGATATCTGATGACGGAGCTTCAATGCTGATCTTTTGACGCATTTGGCTGAG	1283
Db	1330	GTGCTCTGTATGTTTCTGATGATGCTGCTCTTAITCTGGCTGTTCACCTACCCTTGCAGAA	1389
Qy	1284	ACTTCAGAGTTTCTAGGAATGGGTACCATTTTGTGNAGAGTATGACATTTGAACCCAGA	1343
Db	1390	ACATCAGAGTTTCTGCTGAATGGGTACCAATTTTGCAGAGAAATATAGCATAGAGCTTCGT	1449
Qy	1344	GCTCCCGAGTTTCTATTTTCCAGAAATTTGATTACTCTGAAGACAAAGTCCAGCCTTCA	1403
Db	1450	GCACCAAGTGTACTTTGTCTGCAAAATAGATTACTTGAAGGATAAAGTTTCAGACATCA	1509
Qy	1404	TTTGTAAAGACCGCCGGGCGATGAAGAGAGATATGAAGAAATTTAAATTCAGATAAAT	1463
Db	1510	TTTGTCAAAGATCCTAGAGCTATGAAGAGGAAATATGAGGAATTTAAATCCGAATCAAT	1569
Qy	1464	GCCTAGTTTCTAAGCATTTGAAGTCCCGAGGAGGATGATCATGCAAGATGGCACA	1523
Db	1570	GCATTTGTTTCCAAAGCCCTAAATGTCTGAAGAGGGTGGTATGCAAGATGGCACA	1629
Qy	1524	CCATGGCCAGGAACATACAGGGATCATCTCGGAATGATTCAGGTTTTCTTGTGTCAC	1583
Db	1630	CCGTGGCCCTGAAAAATAACAGGGGACCATCCAGGAATGATCCAGGTCCTTTTAGGGCAA	1689
Qy	1584	AGTGTGGCTTCTGATCTAGAGGTATAGAGTCCCGCTTTAGTTATGTCCTCGTGA	1643
Db	1690	AATGGTGGACTTGATGACAGGGCAATGAGTCCCGGTTTGTGTATATGTTTCTCGAGAA	1749
Qy	1644	AAGCTCTCGGTTTCAGACACCAAGAGCTGGTCCCATGAATGCCCTTGTTCGTGTC	1703
Db	1750	AAGGACCAAGATTCAGACACCAAAAGAGCTGGTGTCTAAGATGACATGGTGAGATT	1809
Qy	1704	TCAGCTGTCTCTACTAATGGACAATACATGTTGAATCTTGATTTGATCTACTACATCAAC	1763
Db	1810	TCAGCAGTTCTTACCAATGACCTTTTCATCTTGAATCTTGATTTGATCATTAATAAT	1869
Qy	1764	AACAGCAAGCTGTCCGAGAGCTATGTCCTCTAATGATCCAAACCTAGGTCCGCAA	1823
Db	1870	AACAGCAAAAGCTTAAAGAGAGCAATGTGCTTCTGATGGACCCAAACCTCGGGAAGCAA	1929
Qy	1824	GTCTGTTATGTCAGTTCCACAAAGTTTGTGAGGATTCATAGGAATGATCGATATGCA	1883
Db	1930	GTTTGTTATGTTTCAGTTCCCAAAAGATTGATGTTATCGATGAAGACGATAGATATGCT	1989
Qy	1884	AACAGGAACACTGTCTTTTGTATTAATTAACCTTGAGGGGCTTGACGCCATTCGAAGACCA	1943
Db	1990	AATCGTAATACCGTGTCTTTGTATTAATTAACCTTGAGAGGTTTGTAGATGGATTCGAAGACCT	2049
Qy	1944	GTTTATGTGGAACTGTGTGTGTTTCAACAGAACAGTATCTATGTTTATGAGCCCA	2003
Db	2050	GTATATGTGGAACTGTGATGTGTTTCAACAGAACAGCATTAACGGTTTATGAACCTTCCA	2109
Qy	2004	ATTAAGG-----CGAAGAGACCAAGGTTTCTTGGCATCACTATGTGGGGGCAAGAAG	2057
Db	2110	ATAAAAGTAAACACAGAGACCAAGTCTTTTATCTAAGCTCTGTGTGATCAAGAAAG	2169
Qy	2058	GCAAGCAAGTCAAGAAAAAGAGCTCAGATAAGAAAAAGTCCGAACAAGCATGTGGACAGT	2117
Db	2170	AAGAAATCCAAAGCTAAGAAAGAGTCCGACAAAAGAAATCAGGCAAGGCATACTGACTCA	2229
Qy	2118	TCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGTTTGAAGGTGCTGGGTTTGTAT	2177
Db	2230	ACTGTTCTCTGTTTCAACCTCGATGACATAGAAGAGGAGTTGAAGGTGCTGGTTTGTAT	2289
Qy	2178	GATGAGAAATCAGTTCTCATGTCTCAATCTAGAGAAGAGATTTGGCCAGTCAGCA	2237
Db	2290	GATGAAAAAGCGCTCTTAATGTGCAAAATGAGCTCGAGAGAGGATTTGGACAGTCTGCT	2349
Qy	2238	GCATTTGTTGCCCTCCACTCTGATGGAATATGTTGGTGTCTCAGTCCCTCACTCCAGAA	2297
Db	2350	GTTTTGTGTTCTTACCTTAATGGAATAATGGTGGTGTCTCTCTCTCAGCAACTCCAGAA	2409
Qy	2298	TCTCTTTTGAAGAAGCTATCCATGTCATAAGTTTGGCTATGAGACAGATCTGATGG	2357
Db	2410	AACTTTCTCAAGAGAGCTATCCATGTCAATAGTTTGGTGTATGAGGATAAGTCAGATGG	2469
Qy	2358	GGAACTGAGATTTGGTTGGATCTATGGATCTGTCAAGAAAGATATCTTACTGATTAACAAG	2417
Db	2470	GGAACTGAGATTTGGATGGATCTATGGTCTGTGACAGAAAGATATCTGACTGGTTCAAA	2529
Qy	2418	ATGACGCAAGAGGCTGGGCTTCAGTCTATTTGATGCCCAAGCCGAGCTTTCAAGGGA	2477
Db	2530	ATGATGCGCCGTGGATGGGATCCATTTACTGATGCTTAAGCTTCAGCTTTCAAGGGT	2589
Qy	2478	TCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTCGGTGGGCTCTCGGTTCT	2537
Db	2590	TCTGCTCTCTATCAATCTTTTCAGATCGTCTGAACCAAGTCTGAGGTGGGCTTTAGGTTCA	2649
Qy	2538	GTTGAAATCTTTTTCAGCGGCAATTCGCCCTTATGGTATGGCTACGAGAGGGCGCTCAAG	2597
Db	2650	GTTGAGATCTCTTCAGTCGGCATTCCTATATGATGTTTACAATGGAGGCTTAAAA	2709
Qy	2598	TTCTCTGAGAGATTCGTTTACATCAACACCACTTTACCACTAACCTCTCTCCCGCTT	2657
Db	2710	TTTCTTTCAGAGGTTTCGTATGTGAACACCACTATACCTATCATCTTCTCTCTCT	2769
Qy	2658	CTAGTCTATTGTATATTTGCTGCTATCTGCTCTCACTGGAAGTTTCATCATGCCAGAG	2717
Db	2770	CTCATGTTATGTACATTTGCTAGCCGTTGCTCTTACCAACAGTTTATATTCCTCAG	2829
Qy	2718	ATTAGCAACTTGGCCAGTATCTGGTTCAITTCGCTCTCTCTCTTCAATTTTCGCCACTGGT	2777
Db	2830	ATTAGTAACATTCGAAGTATATGTTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2889
Qy	2778	ATCCTTTCAGATGAGGTGGAGTGGTGTGGCAATGACAGTGGTGGAGGATGAACAGTTC	2837
Db	2890	ATATGAAATAGGTGGAGTGGCTGAGGATAGACGAATTTGGTGGAAACGAGCAGTTT	2949
Qy	2838	TGGGTCAATTCGAGGTATCTCTGCACATCTGTTTGGCGTCTTTCAGGGTCTTCTGAAGGTG	2897
Db	2950	TGGGTCAATTCGAGGTATCCGCTCATTTANTCGTGTGTTTCAAGGTATCTCTCAAGTTC	3009
Qy	2898	CTTGGCGGTATCGACACCACTTCACTGTCCCTCAAGGCTAATGACGAAGAGCGCAC	2957
Db	3010	CTTGGCGGTATTCACACAACTTCACAGTTTACCTCAAAAGCTTCAGATGAAGACGGAGAC	3069
Qy	2958	TTTGTGAGCTCTACATGTTCAAGTGGACGACGCTCTCATCTCTCGACGACCAATTTTG	3017
Db	3070	TTTGTGAGCTCTACTTGTTCAAATGGACAACACTTCTGATTCGCGCAACGACGCTCTC	3129
Qy	3018	ATCATTAACATGTTGTGTGCTTGGCACCTCTTACGCCATCAACAGTGGTTTACCAA	3077
Db	3130	ATTGTAACCTTAGTGGAGTTTGGAGGAGTCTCTTATGCTATCAACAGTGGATACCAA	3189
Qy	3078	TCATGGGGGCGCTCTTTTGGGAAAGCTCTCTTTTGGCTTCTGGGTGATTTGTTCACTTATAC	3137
Db	3190	TCATGGGACCACTCTTTTGGTAAAGTTGTTCTTTTGGCTTCTGGGTGATTTGTTCACTTGTAC	3249
Qy	3138	CCATTCCTCAAGGCTCTTATGGGAGGCAAAAACGACACAGGAGTGTGATCTGCTGG	3197
Db	3250	CTTTCTCAAGGCTTTCATGGGTTCGACGAACCGGACTCTCTCACTTGTGTGGTCTGG	3309
Qy	3198	GCTGTCTCTCTGCTCTATCTTCTCTCTGCTGTGGGTCTGTTGATTCATCACTACAC	3257
Db	3310	TCGTGTTCTTGGCTTCTATCTTCTGTTGTTGGGTAGGATTCATCTCTTCTTCTTCTTCT	3369

[illegible]

Search completed: August 23, 2004, 05:27:48
Job time : 13714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:10:50 ; Search time 1303 Seconds

(without alignments)
11821.923 Million cell updates/sec

Title: US-09-900-237A-29

Perfect score: 3626

Sequence: 1 gcacgagaaacccgcctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002s:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2637.8	72.7	3264	7 ADA69779	Rice gene
2	2595	71.6	3799	6 AAS16458	Corn cDNA
3	2572.2	70.9	3746	3 AAZ299512	DNA encod
4	2572.2	70.9	3746	3 AAZ299527	DNA encod
5	2572.2	70.9	3773	3 AAZ299494	DNA encod
6	2352.4	64.9	3222	7 ADA69499	Rice gene
7	2325.2	64.1	3704	3 AAZ299533	DNA encod
8	1681.6	46.4	1734	3 AAZ58271	Wheat cel
9	1677.6	46.3	3198	3 AAC49550	Arabidops
10	1673.6	46.2	3614	2 AAV06567	Arabidops
11	1408.4	38.8	3725	3 AAZ299500	DNA encod
12	1408.4	38.8	3725	3 AAZ299521	DNA encod
13	1408.4	38.8	3725	3 AAZ299506	DNA encod
14	1382	38.1	3851	3 AAZ299506	DNA encod
15	1373	37.9	3753	3 AAZ299530	DNA encod
16	1373	37.9	3753	3 AAZ299515	DNA encod
17	1373	37.9	3780	3 AAZ299497	DNA encod
18	1343.4	37.0	3603	2 AAV06565	Arabidops
19	1340.2	37.0	3786	3 AAZ58265	Corn cell
20	1340.2	37.0	3813	3 AAZ299509	DNA encod
21	1340.2	37.0	3813	3 AAZ299524	DNA encod
22	1338.6	36.9	3673	2 AAV06568	Arabidops
23	1331.4	36.7	2830	6 AAS16455	Corn cDNA

24	1320.4	36.4	3776	3 AAZ58263	Corn cell
25	1317.2	36.3	3568	3 AAZ299491	DNA encod
26	1271	35.1	3936	3 AAZ58266	Corn cell
27	1271	35.1	3969	3 AAZ299518	DNA encod
28	1271	35.1	3969	3 AAZ299503	DNA encod
29	1257	34.7	3517	3 AAZ58268	Soybean c
30	1217.2	33.6	3328	2 AAV34432	Cotton ce
31	1212.4	33.4	3207	2 AAV08372	Cellulose
32	1201.8	33.1	3747	3 AAZ671145	Eucalyptu
33	1192.6	32.9	3255	6 ABZ12754	Arabidops
34	1191.4	32.9	3255	6 ABZ12754	Arabidops
35	1183.4	32.6	2890	3 AAZ58270	Soybean c
36	1177	32.5	3444	4 AAC83798	Arabidops
37	1176.6	32.4	3828	2 AAV06566	Arabidops
38	1147.2	31.6	3311	2 AAV08373	Cellulose
39	1133.6	31.3	3232	4 AAC65448	Populus t
40	984.2	27.1	2248	2 AAV06562	Arabidops
41	966.8	26.7	2125	3 AAZ58269	Soybean c
42	933.6	25.7	2033	2 AAV08381	Cellulose
43	921	25.4	2306	2 AAV34433	Cotton ce
44	863.4	23.8	1401	3 AAC44966	Arabidops
45	835.4	23.0	2055	3 AAA67111	Eucalyptu

ALIGNMENTS

RESULT 1

ADA69779
ID ADA69779 standard; DNA; 3264 BP.
XX
AC ADA69779;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 3102.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
PS Claim 6; SEQ ID NO 3102; 899pp; English.
XX

The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to

SQ	Sequence	3799 BP; 925 A; 931 C; 995 G; 944 T; 0 U; 4 Other;	
Query Match	71.8%; Score 2595; DB 6; Length 3799;		
Best Local Similarity	87.1%; Pred. No. 0;		
Matches 2871; Conservative	0; Mismatches 420; Indels 4; Gaps 2;		
QY	36 CGGGTTGGATCGCTCTGCGCGCATGGACGGCGAGCGCGACGCCCTGGAAGTCCGGGAGG	95	1056 GAGACCTACCTTGATAGACTGGCTTTAAGTATGACCGAGAAGGTGAACCGTCTCAGTTG
DB	214 CGGTCTGGCGAGCTCGCTCGCATGGAGGGGACCGGACGGGTGAAGTCGGGAGG	273	1234 GAGACTTACCTTGATAGACTCGCATTAAGTATGACCGGAAGGTGAGCCATCTCAGTTG
QY	96 CAGGGGCCGGGACGCTGTCAGATCTGCGCGACCGGCCCTGGGCACACCGTTGGACGC	155	1116 GCTGCTGTTGACATATTTGTGACGTACAGTCGACCCCTTGAAGGAGCCACTATGTCAC
DB	274 CGCGGGGAGGGCAGGTGTCCAGATCTGCGGGATGGCGTGGSCACTACGCGGGAGGA	333	1294 GCTGCTGTTGACATTTTGTGACGTACTGTGCGACCCCAATGAAGGAGCCTCTCTGTG
QY	156 GAGCTTTACCGCTCGACGCTGCGGCTTCCCGGTCTGCGGCTCTGCGGCCCTGCTAG	215	1176 GCCAACCTGTGCTATCCATTCTTGTGTTGATATCCCGTGGACAAAGTCTCTTGTAT
DB	334 GAGCTTTACCGCTCGACGCTGCGGCTTCCCGGTCTGCGGCTCTGCTAGAGTAC	393	1354 GCCAATACCTGTGCTATCCATTCTCGCTGTGGACTATCTGTGATGAAGTCTCTGTG
QY	216 GAGCGCAAGGAGGACCCAGGCTGCTCCAGTGAAGACCAAGTACAAAGCGCCACAGA	275	1236 GTATCTGATGACGAGGCTTCAATGCTGACTTTTGAAGCATTTGGCTGAGACTTCAGAG
DB	394 GAGCGCAAGGAGGACCAAGCTGCCCCAGTGCBAACAAAGTACAAAGCGCCACAG	453	1414 GTATCTGATGAGGCTGCTATGCTGACATTTGATGACATTTGATGACATTTGATG
QY	276 GGGAGCCACCGATCCCGGGGAGGAGGACGACACTGATGCCGATGATGTTAGTGAC	335	1296 GCTAGGAAATGGGTACCATTTGTGAAGAAAGTATGACATTTGAACCGAGCTCCGAG
DB	454 GGGAGTCCAGCATCCGAGGGAGGAGGAGACGATACTGATGCCGATGATGCTAGCG	513	1474 GCTAGAAATGGGTGCCATTTGTTAAGAAGTACAAACATTTGAACCTAGAGCTCCTG
QY	336 TTCAACTTACCTGATCTGACATGAGACGACGACGAGGATGCTGACGAGTCCG	395	1356 TACTTTTGCAGAAATTTGATTTACCTGAAAGACAAAGTCCAGCTTTCATTTGTTAA
DB	514 TTCAACTTACCTGATCTGCAATGACGACGAGGAGGAGGATGCTGACGAGTCCG	573	1534 TACTTCTCCAGAAATTTGATTTGAAGGACAAAGTGCACCTTTCATTTGTTAAAG
QY	396 AGCTGGCGCATGAACACCGGGGAGTGGCAATTTGGCCACCCCAAGTATGACAGTGG	455	1416 CGCGGGCCATGAAGAGAGAAATATGAAGAAATTTAAATCAGGATAAATGCCCTAG
DB	574 AGCTGGCGCATGAATGCTGGGGGACGCGGGATGTTGGCCGCCCAAGTATGACAG	633	1594 CGCGGGCCATGAAGAGAGAAATATGAAGAAATTTAAATCAGGATAAATGCCCTAG
QY	456 GAGATCGGCTCTCAAGATATGACAGTGGAGAGATCCCTAGGGGATACGTCCTTCA	515	1476 AAGGCATTTGAAGTCCCGGAGGAGGATGATCATGCAAGATGGCACACATGGCCAG
DB	634 GAGATCGGCTTACCAAGTACGACAGTGGTGGATCCCTCGGGGATACATCCCGTCA	693	1654 AAGGCACAAAGTCCCTGAGGAAGGATGATCATGCAAGATGGCACACATGGCCAG
QY	516 ACCAAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGGCTGATCATCATGATGTC	575	1536 AACATACAGGATCATCTCGAATGATTCAGGTTTCCCTGTCACAGTGGTGGCTT
DB	694 ACTAAGCCAGATTTCCGGAGAAATCCCTGGTCTTCCCTGACCATCATATGATGTC	753	1714 AACATACAGGACCATCTCGAATGATTCAGGTTTCCCTGTCACAGTGGTGGCTT
QY	576 CCTACGGGAAACATCAGCAGACGCTGCTCCGTTTCCCTATGTAATCATCACAAATCC	635	1596 GATACTGAGGTAATGAGCTCCCGCTTTAGTTTATGTGTCGTGTAAGAGCGTCCCTGG
DB	754 CCTACTGGACATTTGCGAGCGCGCCCAATTTCCCTATATGATCATCAATCCG	813	1774 GATACTGAGGTAATGAGCTACCCGCTTTGCTATGTTTCTCGTGAAGAGCTTCTG
QY	636 TCAAGGAGTTCTCCGCACTATTTGGAAATTTGCGCTGGAAAGAGAGATGATGGCTG	695	1656 TTCAGCACCAAGAGGCTGGTGCATGAATGCGCCCTGTTGCTGTCAGTGTGCTT
DB	814 TCGAGGAAATTTCTCGGTAGCGTTGGGAATGTTGCTGGAAAGAGAGGGTTGATGG	873	1834 TTCAGCATCACAAAGAGCTGGTGCATGAATGCTCTTGTCCGCGCTCTCAGCTGTG
QY	696 ARAATGACAGACAGGGTGCATTTCCCATGACTAATGGCAAGCATTTGCTCCCTCT	755	1716 ACTAATGGACAAATACATGTTGAATCTTGATTTGATCACTACATCAACAACAGAGCT
DB	874 ARAATGACAGACAGGGAAACAAATTTCCCATGCAATGGAATGGCAAGCATTTGCTCC	933	1894 ACCAATGGACAAATACATGTTGAATCTTGATTTGATCACTACATCAACAACAGAGCT
QY	756 GAAGTCCGGCAGCTACTGATCATGATGCTACTGATGATGATGATGATGATGATGATG	815	1776 GTCCGAGAAGCTATGTCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTG
DB	934 GAGGCGGGGTGTTGTTGATATGATGATGATGATGATGATGATGATGATGATGATG	993	1954 CTGAGGAGCTATGTCTTCTTAATGGATCCCTTAACCTAGGAGGAGTGTCTGCTATGT
QY	816 CTGAATGATGAACCTCCGAGCTCTATCTAGAAAGTCCCAATGCTTCTTCCAAATA	875	1836 CAGTTCCTCAAAAGGTTTGAATGAGTATGAGTATGATGATGATGATGATGATGATG
DB	994 TTAACGATGAACCTCCGAGCTCTATCTAGGAAAGTTCACCTTCTTCTCCAGGATA	1053	2014 CAGTTTCCAGAGGTTTCGATGTTATGATAGGAATGATGATATGATGATGATGATG
QY	876 AATCCCTACAGATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	935	1896 GTCTTTTTCATATTAACCTTGAAGGCTTGAAGGCTTCAAGGACAGTTTATGTGGGA
DB	1054 AATCCATACAGATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1113	2074 GTGTTTTTCATATTAACCTTGAAGGCTTGAAGGCTTCAAGGACAGTTTATGTGGC
QY	936 CGTCTCAAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	995	1956 ACTGTTTGTGTTTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAG
DB	1114 CGGATCAAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1173	2134 ACTGGCTGTGTTTTCAACAGAACAGCTATATGTTTATGAGCCCCCAATTAAG
QY	996 ATTTGGTTGCTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1055	2016 AAGCCAGGTTCTTGGCATCATGATGTTGGGGGAGGAGGAGGAGGAGGAGGAGGAG
DB	1174 ATCTGGTTTGTCTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1233	2194 AAGGGTGGTTTCTTGTCTATCATCTATGTTGGGAGGAGGAGGAGGAGGAGGAGGAG
			2252
			2135
			2310

XX	Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;		
SQ	Query Match 70.9%; Score 2572.2; DB 3; Length 3746; Best Local Similarity 86.4%; Pred. No. 0; Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;		
QY	4 CGAGGAACCCCGCTCCAGCTCTCTCGTCCGTCGGGTTGGATCGCTCGCCGCGCATGG 63	Db	1219 ACCATTATGGCTTCTATCTGTTATATGTGAGATCTGGTTTGCTCTTCGTGATATTGG 1278
Db	265 CCAGGCCCAAGGCTCCAGGCCAGCTCCCTCGACGTTTCTCGCGAGCTGCTTGCCATGG 324	QY	1024 ATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTACCTTGATAGATCGCTTTAA 1083
QY	64 ACGGCGACGGACGGCCCTGAAGTCCGGGAGGACACGGGCGGGGACGTGTGCCAGATCT 123	Db	1279 ATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGAGACGTACCTTGATAGCTGGCATTA 1338
Db	325 AGGGCGACGGACGGCGTGAAGTCCGGGAGGCGCGTGGCGACAGGTGTGCCAGATCT 384	QY	1084 GGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTGTGACGTACAG 1143
QY	124 GCGCCACGGCTGGGACACAGTGTGGAACGGACAGTCTTTCACCGCTCGGACGTGTGCC 183	Db	1339 GGTATGACCGGAGAGGTGAGCCATCTCAGTTGCTGCTGTGTGACATTTTGTGACGTACAG 1398
Db	385 GCGGCGACGGCGTGGGACCCACGGCGAGGGGACGTCTTCGCGGCTGCGACGTCTGCG 444	QY	1144 TGACCCCTTGAAGGAGCCACTATCGTCACTGCCAACACTGTGCTATCAATTTCTTGCTG 1203
QY	184 GCTTCCGGTCTCCGCGCCCTGCTAGACGACGAGCGCAAGGAGGCAACCGAGCTGCC 243	Db	1399 TCGACCAATGAAGGAGCCTCCTTGTCACTGCCAATACCGTGTCTATCAATTTCTTGCTG 1458
Db	445 GGTTCGGGTGTCGCGCCCTGCTACGAGTACAGGCGCAGGACGCGACGCGCGTGCC 504	QY	1204 TTGATTATCCCGTGGACAGGTCTCTGTCTATGTATCTGATGACGAGCTTCAATGCTGA 1263
QY	244 TCAGTGCAGAACCAAGTACAAGCGCCACAGAGGAGCCACGATCCGCGGGAGGAAG 303	Db	1459 TGGATTACCTCTGGGATAAGGTCTCTTGTGTATGTATCTGATGATGAGCTGCGATGCTGA 1518
Db	505 CCCAGTGCAGAACCAAGTACAAGCGCCACAAGGGAGCGCGCGATCCGTTGGGAGGAAG 564	QY	1264 CTTTTCACGCATTTGGCTGAGACTTTCAGAGTTTCTAGGAAATGGGTACCATTTGTGAAGA 1323
QY	304 GCGACACACTGATCGGATGATGGTAGTGACTTCAACTACCTGTGATCTGGCACTGAGG 363	Db	1519 CATTTGATGCACCTAGCTGAGACTTTCAGAGTTTGTCTAGAAAATGGGTACCATTTGTGAAGA 1578
Db	565 GAGACACACTGATGCCGA-----TAGCGACTTCAATTACCTTGTGATCTTGGCAATGAGG 618	QY	1324 AGTATGACATTTGAACCCAGAGCTCCGAGTCTTTTACTTTTGCAGAAAAATGTATTACTGA 1383
QY	364 ACCAGAGCAAGAGATTGCTGACAGGATGCGAGCTGGCGCATGAACACCGGGGCGAGTG 423	Db	1579 AGTACAACTTGAACCTAGAGCTCCTGAATGGTACTTCTCCAGAAAAATGTATTACTGA 1638
Db	619 ACCAGAGCAGAGATTGCCGACAGATGCGCAGCTGGCGCATGAACGTTGGGGGACGG 678	QY	1384 AAGACAAAGTCCAGCCTTCATTTGTTAAAGACCGCGGCGCATGAAGAGAGATATGAAG 1443
QY	424 GCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGSCCTCTCCAAGTATGACAGTG 483	Db	1639 AGGCAAAAGTGCAACCTTCAATTTGTTAAAGACCGCGGCGCATGAAGAGAGATATGAAG 1698
Db	679 GGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGACAGTG 738	QY	1444 AATTTAAATTCAGGATAAAATGCCCTAGTTTCTTAAGGCATTTGAAAGTCCCGAGGAAGGAT 1503
QY	484 GAGAGATCCCTAGGGATAGTCCCTTCACTACCAACAGCCAGATGTCAGGAGAAATCC 543	Db	1699 AATTCAAAGTAGGTAAATGGCTTGTGCTAAGSCACAGAAAGTTCTCTGAGAGGAT 1758
Db	739 GCGAGATTCCTCGGGGATACATCCCATGTCATAACAGCCAGATCTCAGGAGAAATCC 798	QY	1504 GGATATGCAAGATGCGACACCATGCGCAGGAAACAAATACCAGGATTCATCTCGGAATGA 1563
QY	544 CTGGAGCTTCGCTGATCATCATGATGTCCTTACGGGAAACATCAGCAGACGTGCTC 603	Db	1759 GGATCATGCAAGATGCGACACCATGCGCAGGAAACAAATACCMGGGACCATCTCTGGAATGA 1818
Db	799 CTGGTCTTCCCTGACCATCATATGATGTCTCCCACTGGAAACATTTGGCAAGCTGCTC 858	QY	1564 TTCAGTCTTTCCTTGGTCAAGTGGTGGCTTGTATCTGAGGTATGAGCTCCCGCTT 1623
QY	604 CGTTTCCCTATGGAATCATTCACAAATCCGTCAGAGGAGTTCTCCGGCAGTATTGGGA 663	Db	1819 TTCAGTCTTTCCTTGGTCAAGTGGTGGCTTGTATCTGAGGCAATGAGCTACCCCGTT 1878
Db	859 CATTTCCCTATGTGAACCAATTCGCCAAATCCGTCAGGGAGTTCTCTGGTAGCATTTGGGA 918	QY	1624 TAGTTTATGTGTCCTGTAAGAGCGTCTGGGTTCAGGACCCACAGAAAGGCTGGTGCCA 1683
QY	664 ATGTTCCCTGGAAGAGAGTTGATGGCTGGAATGAGCAGGACAGGGTGGCATTC 723	Db	1879 TGCTCTATGTTTCTCGTGAAGAGCGTCTTGGATTCAGCATCACAGAAAGCTGGTGCCA 1938
Db	919 ATGTTCCCTGGAAGAGAGGTTGATGGCTGGAATGAGCAGGACAGGGGACGATTC 978	QY	1684 TGAATGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1743
QY	724 CCATGACTAATGGGACAAAGCATTCCTCTCAAGGTCCGGCAGCTACTGACATCGATG 783	Db	1939 TGAATGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1998
Db	979 CCATGACGAATGCAACAGCATTTGCTCCCTCTGAGGGTCCGGGTGTTGGTGATATTGAIG 1038	QY	1744 ATTGTGATCACTACATCAACACAGCAAGGCTGTCGGAAGCTATGTGCTTCTCTAATGG 1803
QY	784 CATCTACTGAATACACATGGAAGAGCGTTTACTGAATGATGAAATCCGCGACGCTCTAT 843	Db	1999 ATTGTGATCACTACATCAACACAGTAAGGCTCTCAGGAAAGCTATGTGCTTCTCTAATGG 2058
Db	1039 CATCAACTGATTACACATGGAAGATGCTTATTGAACGACGAAATCTCGACAGCCTCTAT 1098	QY	1804 ATCCAAACCTAGTCCGCAAGTCTGTTATGTGAGTTCACCAAAAGGTTTGTATGGGATTG 1863
QY	844 CTAGAAAAGTCCCATTTGCTCTCCAAATAAATCCCTACAGAATGGTCATTGTTCTGC 903	Db	2059 ACCCTAACCTAGGAAGAGTGTCTGCTACGTCCAGTTTCCCGAGATTCGATGGCATTTG 2118
Db	1099 CTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGATGGTCAITTTGCTGTC 1158	QY	1864 ATAGGAATGATCGATNGCAACAGGAACACTGTCTTTTGTATTAATTAATCTGAGGGGCC 1923
QY	904 GGTGTTGTTCTAAGCATCTTCTGCACTACCGTCTCAAAATCCGTGCGGTAATGCAAT 963	Db	2119 ACAGGAATGATCGATNGCAACAGGAACACCGTGTCTTTCGATATTAATCTTGAAGGTC 2178
Db	1159 GATTGATTGTTCTAAGCATCTTCTTGCATCTACCGTATCAAAATCCGTGCGCAATGCAAT 1218	QY	1924 TTGACGGCATCAAGGACCAAGTTTATGTGGGAACCTGGTTGTGTTTTTCAACAGAACAGCTA 1983
QY	964 ACCCACTGGGCTTTTATCTGTTATATGTGAGATTGTTGTTTGTCTTTATCTGATACTGG 1023	Db	2179 TTGATGGCATCCAGGACCAAGTTTATGTGGAACTGGCTGTGTTTCAACCCGAAACAGCTC 2238
		QY	1984 TCTATGGTTATGAGCCCCCAATTAAAGCGAAGACCAAGCTTCTTTGGGATCACTATGTG 2043
		Db	2239 TATATGTTTATGAGCCCCCAATTAAAGCAGAAAGGGTGGTTTCTTGTCTATCACTATGTG 2298
		QY	2044 GGGGCAAGAGAGGCAAGCAAGTCAAGAAAGGAGCTCAGATAGAAAGAAAGTCCGAACA 2103
		Db	2299 GCGGTAGGAAGAGGCAAGCAAAATCAAGAA---GGGCTCGGACAAAGAAAGTTCGAGA 2355

QY	2104	AGCATGTGGACAGTTCCTGTTCCAGTATTCAATCTCGAAGACATAGACGAGGGTGTTCCAAG	2163
Db	2356	AGCATGTGGACAGTTCCTGTTGCCAGTATTCAACTTGAAGATATATAGACGAGGGAGTTCCAAG	2415
QY	2164	GTCCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCCTCAAATGAGCTTAGAGAAGAGAT	2223
Db	2416	GCGCTGGATTGACGACGAGNAATCACCTTCTTAATGCTCAANYGAGCCTCGAGANAAGAT	2475
QY	2224	TTGCCCACTCAGCASCATTTTGTGCCTCCACTCTGATGGAATATGGTGGTGTTCCTCAGT	2283
Db	2476	TTGCCCACTCGCAGCGCTTGTGTGCTCCACTCTGATGGAGTATGGTGGTGTTCCTCAGT	2535
QY	2284	CCTCACTCCAGAAATCTCTTTGAAAAGAGCTATCCATGTCAATGATTTGGCTATGAGG	2343
Db	2536	CCGCAACTCCGGAGTCTCTTCTGAAAGAAAGTATCCATGTTATAAGCTGTGGCTATGAGG	2595
QY	2344	ACAAGTCTGAATGGGAACCTGAGANTCGTGGATCTATGGATCTGTACACRAGATATTC	2403
Db	2596	ACAAGTCTGAATGGGAACCTGAGATCGGTGGATCTACGGTCTCTGTACACRAGAACATTC	2655
QY	2404	TTACTGGATTCAAGATGCACGAAAGAGCTGGCGCTTCAGTCTATTGCATGCCAACGCGCC	2463
Db	2656	TCACCGGATTCAAGATGCACGCGAGAGCTGGCGGTGCACTCTACTGATGCCAAGCGGC	2715
QY	2464	CAGCTTTCAAGGGATCTGCCCCCCATCAATCTTTTCAGATCGTCTGAAACCAAGTGCTCGCGT	2523
Db	2716	CAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGGACCGTCTGAAACCAAGTGCTCGCGT	2775
QY	2524	GGGCTCTCGGTTCTGTTGAAATCTTTTTTCAGCGGCAATGGCCCTTATGTTATGGCTACG	2583
Db	2776	GGGCTCTTGGGTCCGTGGAGATCCCTTCACGCGGCACTGCCCTGTGTTACGGCTACG	2835
QY	2584	GAGGGCGCCTCAAGTTCCTCGAGAGATTGCTGTACATCAACACCACCATTTACCCACTAA	2643
Db	2836	GAGGGCGGCTCAAGTTCCTGGAGNATTCGGGTACATCAACACCAATCTACCCGCTCA	2895
QY	2644	CCTCTCTCCCGTPTCTAGTCTATTGTATATGTCCTGCTATCTGTCTCTCACTGGAAAGT	2703
Db	2896	CGTCCATCCCGCTCTCATCTACTGTCATCTGTCGCCGCCATCTGTCTCTCACCGAAAGT	2955
QY	2704	TCATATGCACAGATTTAGCAACTTGGCCAGTATCTGGTTCATTTGCGCTCTTCTTTCAA	2763
Db	2956	TCATCATTCACAGATCAGCAACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCATCTCGA	3015
QY	2764	TTTTTCGCCACTGGTATCTCTTTGAGATGAGTGGAGTGGTGTGGCATTTGACGAGTGGTGA	2823
Db	3016	TCCTTCGCCACGGGCATCCTCGAGATGAGTGGAGCGGGGTGGGCATTCGACGAGTGGTGA	3075
QY	2824	GGAAATGAACAGTTCTTGCGTCAITGGAGGTATCTCTGCACATCTGTTTGGCGCTCTTTCAGG	2883
Db	3076	GGAAACGAGCAGTTCTTGCGTCAITGGGGGCATCTCCGGCGACCTCTTCGCGGTTCGAGG	3135
QY	2884	GTCTTCTGAGAGTGTCTTGCCGGTATCGACACCAACTTTCATCTGTCACCTCAAGAGCTAATG	2943
Db	3136	GCCTGTCTCAAGTGTCTGCGCGGCATCGACACCAACTTTCAGCGTCACTCCCAAGGCGCTCGG	3195
QY	2944	ACGAAAGGCGCATTTWGTCTGAGCTTACATGTTTCAAGTGGAGCAGCGCTTCTCATCCCTC	3003
Db	3196	ACGAGHACGGCGACTTTCGCGAGCTGTACATGTTTCAAGTGGAGCAGCTCTCTGATCCCGC	3255
QY	3004	CGACGACCATTTTGATCATTTAAACATGGTGTGGTGTGCTGTCGACCTCTCTAGCCCATCA	3063
Db	3256	CCACCAACATCTGATCATCAACTGTGTGGCGTCTGTCGCGGCATCTCTTAGCCCATCA	3315
QY	3064	ACAGTGTATTACCAATCATAGGGGGCGCTCTTTTGGGAAAGCTCTTCTTTTGCCTTCTGGGTGA	3123
Db	3316	ACAGCGGATACCAAGTCTGTGGGGCCGCTCTTTCGGCAAGCTCTTCTTCGCTCTCTGGGTCA	3375
QY	3124	TTGTTCACTTATACCAATCTCTCAAGGTTCTTATGGCAGGCAAAACCGCACACCGACGA	3183
Db	3376	TCGTTCACCTGTATACCGGTCTCTCAAGGGCTCTATGGCAGGCGAAACCGCACCCCGACCA	3435

QY	3184	TTGTCATCGCTGGCGTGTCTCTCGCTTCTATCTTCTCTCTGCTGGTTCGTGTTG	324	
Db	3436	TCGTGCTGCTGTGGGCATCTCTGCTGGCTCCATCTTCTCTCTGCTGGTTCGCATCG	3495	
QY	3244	ATCCATTCACTACCCGCTCTCGCTGGCCCAATAATCCAAACCTGTGGCATCAACTGCTAGG	3303	
Db	3496	ACCCCTTCACCAACCGCGTCACTGCGCCGGATACCCAGAGCTGTGGCATCAACTGCTAGG	3555	
QY	3304	AAAGTGGGAGTTT	3316	
Db	3556	GAAGTGAAGGTT	3568	
RESULT 4				
ID	AAZ99527	standard; DNA; 3746 BP.		
XX	AC	AAZ99527;		
DT	03-JUL-2000	(first entry)		
XX	DNA encoding a maize cellulose synthase.			
XX	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;			
KW	transgenic plant; plant breeding marker; ss.			
XX	Zea mays.			
OS	Key	Location/Qualifiers		
FH	CDS	321...3449		
FT		/*tag= a		
FT		/product= "cellulose synthase"		
FT		/transl_except= (pos: 1800..1802, aa: Xaa)		
FT		/note= "no termination codon given; Xaa is an unspecified		
FT		amino acid"		
XX				
PN	WO200009706-A2.			
XX				
PD	24-FEB-2000.			
XX				
PF	16-AUG-1999;	99WO-US018760.		
XX				
PR	17-AUG-1998;	98US-0096822P.		
XX				
PA	(PTON-)	PIONEER HI-BRED INT INC.		
XX				
PI	Dhugga KS, Helentjaris TG, Bowen BA,	Wang X;		
XX				
DR	WPI; 2000-224343/19.			
DR	P-PSDB; AAY84119.			
XX				
PT	New genes which encode maize cellulose synthase polypeptides in plants			
PT	useful for modulating the expression of cellulose synthase in plants and			
PT	to produce transgenic plants expressing the novel protein.			
XX				
PS	Claim 1; Page 176-181; 119pp; English.			

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of

CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ

Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

Query Match 70.9%; Score 2572.2; DB 3; Length 3746;

Best Local Similarity 86.4%; Pred. No. 0;

Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

QY 4 CGAGGAACCCGCTCCAGCTCTCTCGTTCGGTGGGTGGATCGCTCTGCCGCGCATGG 63

DB 265 CAGGGCCCAAGGCTCCAGGCCAGCTCCCTCGAGTCTCTCGGCGAGCTCGCTGGCCATGG 324

QY 64 ACGGCGACCGGACGGCCCTGAAGTCCGGGAGGACCGGGCCGGGAGCGTGTGCCAGATCT 123

DB 325 AGGGCGACCGGACGGCGTGAAGTCCGGGAGGCGCGGTGGCGGACAGGTGTGCCAGATCT 384

QY 124 GCGCCACGGCTGGGACACACGCTTGGACGGGACGCTCTTCACGGCTCGAGCTGCC 193

DB 385 GCGCGACCGGCTGGGACACACGCGGAGGGGACGCTCTTCGCGGCTTCGACGCTGGCG 444

QY 184 GCTTCCGGTCTCGCGCCCTGCTACGACGACGAGCGCAAGGAGGCGCACCGGCTGCC 243

DB 445 GGTTCGCGTGTGCGGCCCTGCTACGAGTACGAGTACGAGCGCGAGGACGCGCGAGCTGCC 504

QY 244 TCCAGTGCAGAACCAAGTACAAGCGCCACAGAGGGAGCCCGAGATCCGCGGGAGGAAG 303

DB 505 CCCAGTGCAGAACCAAGTACAAGCGCCACAAAGGGAGCGCGGCGATCCGCTGGGAGGAAG 564

QY 304 GCGACGACCTGATCGGATGATGGTAGTCACTTCACTACCTCGATCGGCACTGAGG 363

DB 565 GAGACGACCTGATGCCGA-----TAGCGACTTCAATTACCTTGTCACTTGGCAATGAGG 618

QY 364 ACCAGAAGCAGAGATGCTGACAGGATGCGCAGCTGGCGCATGAACACCGGGGCGAGT 423

DB 619 ACCAGAAGCAGAGATGCGCAGAGATGCGCAGCTGGCGCATGAACGTTGGGGGAGCG 678

QY 424 GCAATGTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACAGT 483

DB 679 GGGATGTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGACAGT 738

QY 484 GAGAGATCCCTAGGGATAGTCCCTCACTACCAACAGCGAGATGTCAGGAGAAATCC 543

DB 739 GCGAGATTCCTCGGGGATACATCCCATCACTCAACAGCCAGATCTCAGGAGAAATCC 798

QY 544 CTGGAGCTTCGCTGATCATCATGATGTCCCTACGGGGAACATCAGCAGACGTGCTC 603

DB 799 CTGGTCTTCCTGACCATCATATGATGTCCCACTGGGGAACATTTGGCAAGCGTCTC 858

QY 604 CGTTCCCTATGTAATCATTCACCAATCCGTCAGGGAGTCTCCGGCAGTATTGGGA 663

DB 859 CAITTCCTATGTGAACCATTCGCCAAATCCGTCAAGGGAGTCTCTGTGTAGCATTTGGGA 918

QY 664 ATCTTCCTGGAAGCAGAGTGTGATGGCTGGAATGAGCAGGACAAAGGTGCGATTC 723

DB 919 ATGTTCCTGGAAGCAGAGTGTGATGGCTGGAATGAGCAGGACAAAGGGACGATTC 978

QY 724 CCATGACTAATGGGACAAAGCATGTCTCCCTCTGAAGGTCCGGCAGCTACTGACATCGATG 783

DB 979 CCATGAGATGGCAGCAGCATGTCTCTCTGAGGCTCGGGGTGTGGTATATTGATG 1038

QY 784 CATCTACTGAATCAACATCGAAGACGCTTTACTGAATGATGAACATCGCCAGCTCTAT 843

DB 1039 CATCAACTGATTACCAATCGAAGATGCCCTTATTGAACGACGAAATCTCGACAGCTCTAT 1098

QY 844 CTAGAAGATCCCATTTCTCTCCAAAATAAATCCCTACAGAAATGGTCAATTTCTGCG 903

DB 1099 CTAGAAGATTCACATTCCTCTCCAGGATAAATCCATACAGGATGGTCAATTTGCTGCG 1158

QY 904 GGTGGTGTGTTAAGCATCTTCCTGCATACCGTCTCAAAAATCCCTGTGCGTAATGCAT 963

DB 1159 GATTGATGTTCTAAGCAATCTTCTTGCATCTACGTTATCAAAATCTCTGTGCGCAATGCAT 1218

QY 964 ACCACTGTGGCTTTTATCTGTATATGTAGATTTGGTTTGTCTTATCTCTGATGACTGG 1023

DB 1219 ACCCAATTATGGCTTCTATCTGTATATGTAGATCTGGTTTGTCTTCTCTGTGATATTGG 1278

QY 1024 ATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGTATAGACTCGCTTTAA 1083

DB 1279 ATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGTATAGCTGGCAATA 1338

QY 1084 GGTATCAGCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTGTTCAGTACAG 1143

DB 1339 GGTATCAGCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGTGACATATTGTTCAGTACAG 1398

QY 1144 TCGACCCCTTGAAGGAGCCACTATCTGCTCACTGCCCAACTGTGCTATCCATCTTGTCTG 1203

DB 1399 TCGACCAATGAAGGAGCCCTCTCTTGTCACTGCCAATACCGTGTCTATCAATCTTGTCTG 1458

QY 1204 TTGATTATCCCGTGGACAAGTCTCTTGTCTATGTATCTGATGACGAGCTTCAATGCTGA 1263

DB 1459 TGGATTACCTGTGGATAAGTCTCTTGTCTATGTATCTGATGATGAGCTGCGATGCTGA 1518

QY 1264 CTTTTCACGCATTGGCTGAGACTTTCAGAGTTTCTTAGGAAATCGGATCCATTTGTGAAGA 1323

DB 1519 CATTTGATGCATAGCTGAGACTTTCAGAGTTTCTTAGGAAATCGGATCCATTTGTGAAGA 1578

QY 1324 AGTATCACAATTGAACCCAGAGCTCCCGAGTTTCTTTTCCAGAAAATTTGATTAATCTGA 1383

DB 1579 AGTATCAACATTGAACCTGAGACTTCTGAAATGGTACTTCTCCAGAAAATTTGATTAATCTGA 1638

QY 1384 AAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAG 1443

DB 1639 AGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAG 1698

QY 1444 AATTTAAATCAGGATAAATGCCCCTAGTTTCTAAGGCAATTTGAAAGTCCCGAGGAAGGAT 1503

DB 1699 AATTTCAAGTATGGTAAATGGCTTGTGTTAAGGCACAGAAAATTTCTCTGAGGAAGAT 1758

QY 1504 GGATCATGCAGATGACACACCATGGCCAGGAACAATACCAGGGATCATCTCTGGGAATGA 1563

DB 1759 GGATCATGCAAGATGACACACCATGGCCAGGAACAATACCAGGGATCATCTCTGGGAATGA 1818

QY 1564 TTCAGGTTTCTCTGTGTCACAGTGGTGGCTTCTGATCTAGAGGTAATAGCTCCCGCTT 1623

DB 1819 TTCAGGTTTCTCTGTGTCACAGTGGTGGCTTCTGATCTAGAGGCAATAGCTACCCGCTT 1878

QY 1624 TAGTTTATGTGTCCTGTGAAAAGCGTCTGGGTTCCAGCACCAAGAGAGGCTGGTGCCA 1683

DB 1879 TGGTCTATGTTTCTCGTAAAAGCGTCTGGATTCAGGCATCACAAAGAAAGCTGGTGCCA 1938

QY 1684 TGAATGCCCTTGTTCGTGTCCTGCTGCTCTACTAATGAGCAATACATGTTGAAATCTTG 1743

DB 1939 TGAATGCTCTGTTGTCGTGTCCTGCTGCTTACCAATGAGCAATACATGTTGAAATCTTG 1998

QY 1744 ATTGTGATCACTACATCAACACAGCAAGGCTCTCGAGAGGATATGTGCTTCTCTAATGG 1803

DB 1999 ATTGTGATCACTACATTAACACAGTAAGGCTCTCAGGAGAGCTATGTGCTTCTCTATGG 2058

QY 1804 ATCCAAACCTTAGTCCGCAAGTCTGTTATGTGAGTTTCCCAAAAGTTTGTATGGGATG 1863

DB 2059 ACCCTAACCTTAGGAAGGAGTCTGTCTACGTCCAGTTTCCCGAGAGATTCGATGGCATG 2118

QY 1864 ATAGGAATGATCATATGCAACAGGACACTCTCTTTTGTATATTAACTTCAGGGGCT 1923

DB 2119 ACAGGAATGATCATATGCAACAGGACACTCTCTTTTGTATATTAACTTCAGGAGCT 2178

QY 1924 TTCACGGCATTCAGGACCAAGTCTTATGTGGAACTTGTGTGTTTTTCAACAGAACAGCTA 1983

DB 2179 TTGATGGCATTCAGGACCAAGTCTTATGTGGAATCTGCTGTGTTTTTCAACCGAACAGCTC 2238

QY 1984 TCTATGTTATGAGCCCCCAATTAAGCGAGAGGACCGAGTCTTCTTGGCATCACTATGTC 2043

Db 2239 TATATGGTTATGAGCCCCCAATTAAGCAGAAAGAGGGTGGTTTCTTGTCTACCTCATGTG 2298
QY 2044 GGGCAAGAGAGCAGCAAGCTCAAGAAAGAGAGCTCAGATAAGAAAAGTCGACAA 2103
Db 2299 GCGGTAGGAAGAGCAAGCAAAATCAAGAA---GGGCTCGGACAAGAAAGAGTCGAGA 2355
QY 2104 AGCATGTGGACAGTTCTGTCCAGTATTCAATCTCGAAGACATAGAGGAGGGTGTGAAG 2163
Db 2356 AGCATGTGGACAGTTCTGTCCAGTATTCAATCTCGAAGACATAGAGGAGGGTGTGAAG 2415
QY 2164 GTGCTGGTGTGATGATGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAGAAGAGAT 2223
Db 2416 GCGCTGGATTTGACACAGAGAAATCACTTCTTATGTCTCAATGAGCCTGAGAAGAGAT 2475
QY 2224 TTGGCCAGTACAGCAGCATTTGTTGGCTCCACTCTGATGGAATATGGTGTCTCTCAGT 2283
Db 2476 TTGGCCAGTACGCGAGCGTTGTTGGCTCCACTCTGATGGAATATGGTGTCTCTCAGT 2535
QY 2284 CCTCCACTCCAGAAATCTCTTTTGAAGAAAGCTATCCATGTCTAATAGTTGTGGCTATGAG 2343
Db 2536 CCGCAACTCCGAGTCTCTCTGAAGAAAGCTATCCATGTCTAATAGTTGTGGCTATGAG 2595
QY 2344 ACAAGTCTGAATGGGAATCAGATGGTGGATCTATGGAATCTGATGGAATCTGACAGAAATATC 2403
Db 2596 ACAAGACTGAATGGGAATCAGATGGTGGATCTATGGAATCTGATGGAATCTGACAGAAATC 2655
QY 2404 TTACTGGATTCAGATGACACCAAGAGGCTGGGTTCACTTATGATGCAATGCAAGCGCC 2463
Db 2656 TCACCGGATTCAGATGACACCGGAGGCTGGGTTCACTTATGATGCAATGCAAGCGCC 2715
QY 2464 CAGCTTTCAAGGATCTGCCCCCATCAATCTTTGATGCTGAAACCAAGTGTGCGGT 2523
Db 2716 CAGCTTTCAAGGATCTGCCCCCATCAATCTTTGATGCTGAAACCAAGTGTGCGGT 2775
QY 2524 GGGCTCTGGTCTGTGTAATCTTTTACGCGGCAATGCGCCCTATGATGATGCTACG 2583
Db 2776 GGGCTCTGGTCTGTGTAATCTTTTACGCGGCAATGCGCCCTATGATGATGCTACG 2835
QY 2584 GAGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCAATTTACCCACTAA 2643
Db 2836 GAGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCAATTTACCCACTAA 2895
QY 2644 CCTCTCTCCGCTTCTAGTCTATTGATATTGATGCTGCTATCTGCTCACTGCAAGT 2703
Db 2896 CGTCAATCCGCTTCTCATCTGATCTGATCTGCGGCACTGCTGCTCAAGCAAGT 2955
QY 2704 TCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTCAATTCGCTCTTCTTTCAA 2763
Db 2956 TCATCATGCCAGAGATCAGCAACTTGGCCAGTATCTGTTCAATTCGCTCTTCTTTCAA 3015
QY 2764 TTTTCCGCACTGGTATCCTTGAGATGAGTGGAGTGTGTTGGCATTCACGAGTGTGA 2823
Db 3016 TCTTCCGCACTGGTATCCTTGAGATGAGTGGAGCGGGTGGGATTCACGAGTGTGA 3075
QY 2824 GGAATGAACAGTCTTGGGTCAATGGAGTATCTCTGCAATCTGTTTCCGCTCTTTCAGG 2883
Db 3076 GGAACGACAGTCTTGGGTCAATGGAGTATCTCTGCAATCTGTTTCCGCTCTTTCAGG 3135
QY 2884 GTCCTTGAAGGTGCTGCGGATTCGACACAACTTCACTGTCACTCAAGGCTAATG 2943
Db 3136 GCCTGCTCAGGTGCTGCGGATTCGACACAACTTCACTGTCACTCAAGGCTAATG 3195
QY 2944 ACGAAGAGGCACTTTGCTGAGCTCTACATGTTCAAGTGAAGAGCGTCTTCACTCCCTC 3003
Db 3196 ACGAAGAGGCACTTTGCTGAGCTCTACATGTTCAAGTGAAGAGCGTCTTCACTCCCTC 3255
QY 3004 CGACGACCAATTTGATCAATTAACATGTTGTGTTGCTGTTGCTGCACTCTTCAAGGCTA 3063
Db 3256 CCACCAATCTGATCAATCACTGTTGCTGTTGCTGCACTCTTCAAGGCTA 3315
QY 3064 ACAGTGTGTACCAATCATGCGGGCGGCTCTTTGGGAAGCTCTTCTTTCCTTCTGGGTGA 3123
Db 3316 ACAGCGGATACAGTCTGTTGGGCGGCTCTTTCGGAAGCTCTTCTTCTGCTTCTGGGTGA 3375

QY 3124 TTGTTCACTTATACCATTCCTCAAGGGTCTTATGGGAGGCAAAACCGCACCGACGA 3183
Db 3376 TCGTCCACCTGTACCCGTTCTCAAGGGCTCTATGGGAGGCAAAACCGCACCGACGA 3435
QY 3184 TTGTCATCGTCTGGGCTGTCTCTCGCTTCTATCTTCTCTCTGCTGTTGGGTTCGTGTG 3243
Db 3436 TCGTCTGCTGTGGGCAATCTCTGCTGGGTCATCTTCTCTGCTGTTGGGTTCGCATCG 3495
QY 3244 ATCCATTCACTACCGCTCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACTGCTAGG 3303
Db 3496 ACCCTTTCACCAACCGGCTCACTGGCCGAGTACCCAGCTGTGGCATCAACTGCTAGG 3555
QY 3304 AAAGTGGAGTTT 3316
Db 3556 GAAGTGAAGTTT 3568

RESULT 5
AAZ99494
ID AAZ99494 standard; DNA; 3773 BP.
XX
AC AAZ99494;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 337..3565
FT a /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= [pos: 1817..1819, aa: Xaa]
FT /note= "no termination codon given; Xaa is an unspecified amino acid"
XX
PN MO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
XX
DR P-PSDB; AAY84108.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 90-94; 119pp; English.
XX

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting

CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX

SQ Sequence 3773 BP; 863 A; 988 C; 1017 G; 904 T; 0 U; 1 Other;

Query Match 70.9%; Score 2572.2; DB 3; Length 3773;

Best Local Similarity 86.4%; Pred. No. 0;

Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

QY	4	CGAGGAACCCGCTCCAGCTCTGTGCTGCGGTGGATGCTCTGCGCGCATCG 63
DB	282	CCAGGCCAGGCTCCAGGCCAGCTCCCTCGACGTTTCTCGGCGAGCTGCTTGCCATGG 341
QY	64	ACGGCAACCGGACGCCCTGAAGTCGGGAGGACCGGGCCGGGACGTGTGCCATCT 123
DB	342	AGGGCGACCGGACGGCGTGAAGTCGGGGAGCGCGGTGGCGACAGGTGTGCCATCT 401
QY	124	GCGCCACGCGCTGGGCACACGTTGCGACGGGACGCTTTCACCGCTCGGACGTCTGCC 183
DB	402	GCGCGACGCGCTGGGCACACGTTGCGACGGGAGCGCTTTCGCGCGCTCGGACGTCTGG 461
QY	184	GCTTCCCGTCTGCGCCCTCTGTACGAGCAGAGCGCAAGGAGGACCCAGGCTTGC 243
DB	462	GGTTTCCGCTGTGCGCCCTCTGTACGAGTACGAGCGAAGGACGCGACGCGGCTGCC 521
QY	244	TCCAGTGCAGACCAAGTACAAGCGCCACAGGGGAGCCGAGGATCCGGGGGAGGAAG 303
DB	522	CCCAGTGCAGACCAAGTACAAGCGCCACAGGGGAGCCGCGGATCCGTTGGGGAGGAAG 581
QY	304	GCGACGACACTGATCGCGATGATGTAGTACTTCAACTACCTGCTTGGCACTGAGG 363
DB	582	GAGACGACACTGATGCCGA-----TAGCGACTTCAATTACCTTGGATCTGGCAATGAGG 635
QY	364	ACCAGAAGCAGAAGATTGTGACAGGATCGGACGTGGCGCATGAACACCGGGGCGAGTG 423
DB	636	ACCAGAAGCAGAAGATTGCCGACAGAATCGCGAGCTGGCGCATGAACGTTGGGGGCGAG 695
QY	424	GCATCTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACAGTG 483
DB	696	GGGATGTTGGTGGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGACAGTG 755
QY	484	GAGAGATCCCTAGGGGATAGCTCCCTTCAGTCAACCAAGCCAGCATGTGAGGAGAATCC 543
DB	756	GCGAGATTCTCTGGGATATACCCCATCAGTCACTAACGCCAGATCTCAGGAGAATCC 815
QY	544	CTGGAGCTTGGCTGATCATCATGATGTCCCTTACGGGGAAACATCAGCAGCGTCTC 603
DB	816	CTGGTCTTCCCTGACCATCATATGATGTCCCAACTCGGAACATTTGGCAAGCGTCTC 875
QY	604	CGTTTCCCTATGTGATCATTCACCAATCCGTCAGGGAGTCTCCGGCAGTATTGGGA 663
DB	876	CATTTCCCTATGTGAACCAATCGCCAAATCCGTCAGGGAGTCTCTGGTAGCATTTGGGA 935
QY	664	ATGTTGCTTGGAAAGAGAGAGTTGATGGCTGGAATAATGAAGCAGGACAAGGGTGGCATTC 723
DB	936	ATGTTGCTTGGAAAGAGAGAGTTGATGGCTGGAATAATGAAGCAGGACAAGGGGACCATTC 995
QY	724	CAATGACTAATGGGACAAGATTGCTCCCTCTGAAGGTGCGGCGAGCTACTGACATCGATG 783
DB	996	CCATGACGAATGGACAAGCAATGCTCCCTCTGAGGGTGGGGTGTGGTGATATTGATG 1055
QY	784	CATCTACTGAATACATGGAAGCGCTTACTGAAATGATGAACTCCGCCGCTCTAT 843
DB	1056	CATCAACTGATTACAACTGGAAGATGCCTTATTGAACGACGAAACTCGACACCTCTAT 1115

QY	844	CTAGAAAAGTCCCATTTGCTTCTCCAAAATAAAATCCCTACAGAAATGGTCATTGTTCTGC 903
DB	1116	CTAGAAAAGTTCACATTCCTTCTCCAGGATAAATCCATACAGGATGGTCATTGTGCTGC 1175
QY	904	GGTTGGTTGTTTAAAGCATCTTCTGACATCCGCTCTCACAATCTCTGGGTAATGCAAT 963
DB	1176	GATTGATGTTCTTAAGCATCTTCTGACATCCGATATCACAATCTCTGGGCAATGCAAT 1235
QY	964	ACCCACTGGGCTTTTATCTGTATATGAGATTGGTTGGTTGCTTTATCTCGATGACTGG 1023
DB	1236	ACCCATATGGGTTCTATCTGTATATGAGATCTGGTTTGTCTTTCTCGTGATATGG 1295
QY	1024	ATCAGTTTCCGAAAGTGGTTTCCAATCAACCGGAGACCTACCTTGATAGACTGGCTTTAA 1083
DB	1296	ATCAGTTTCCCTAAGTGGTTTCCAATCAACCGGAGACGTACCTTGATAGCTGGCATTAA 1355
QY	1084	GGTATGACCGGAGAGGTGAACCGTCTCAGTTGGTCTCTGTTGACATATTGTCAGTACAG 1143
DB	1356	GGTATGACCGGAGAGGTGAGCCATCTCAGTTGGTCTCTGTTGACATTTTCTGTCAGTACAG 1415
QY	1144	TCGACCCCTTGAAGGAGCCACCTATGCTCACTGCCAACACTGTGCTATCCATTCTTGTCTG 1203
DB	1416	TCGACCCAAATGAGGAGCGCTCTCTTGTCACTGCCAATACCGTGTCTATCCATTCTTGTCTG 1475
QY	1204	TTGATTATCCCGTGGACAAGGCTCTCTTGTCTATGTATCTGATGACGAGGCTTCAATGCTGA 1263
DB	1476	TGGATTTACCTCTGGATAGGCTCTCTTGTCTATGTATCTGATGATGAGCTGCGATGCTGA 1535
QY	1264	CTTTTGAACCATTTGGCTGAGACTTTCAGAGTTTGTCTAGGAAATGGGTACCATTTGTGAAGA 1323
DB	1536	CATTTGATGCACTAGCTGAGACTTTCAGAGTTTGTCTAGGAAATGGGTACCATTTGTGAAGA 1595
QY	1324	AGTATGACATTGAACCCAGAGCTCCGAGTTTCTCTTTTCCAGAAAAATGATTACTCTGA 1383
DB	1596	AGTACAACATTGAACCTAGAGCTCTGAAATGGTACTTCTCCAGAAAAATGATTACTCTGA 1655
QY	1384	AAGACAAATCCAGCGCTTCAATTTGTTAAAGACCGCGGCCCATGAAGAGAGAAATGAAG 1443
DB	1656	AGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCGGCCCATGAAGAGAGAAATGAAG 1715
QY	1444	AATTTAAATTCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAAGTCCCGAGGAAGGAT 1503
DB	1716	AATTTAAAGTTAGGTTAAATGGCTTGTGTTGTTAAGGCACAGAAAGTCTCTGAGGAAGGAT 1775
QY	1504	GGATCATGCAAGATGSCACCATGSCCAGGAAACAAATACCAGGATCATCTCTGGAATGA 1563
DB	1776	GGATCATGCAAGATGSCACCATGSCCAGGAAACAAATACCAGGATCATCTCTGGAATGA 1835
QY	1564	TTCAGGTTTTCTTGGTCAAGTGGTGGCTTTGATACCTAGGGTAATGAGCTCCCCCGTT 1623
DB	1836	TTTCAAGTTTTCTTGGTCAAGTGGTGGCTTTGATACCTAGGGCAATGAGTACCCCGTT 1895
QY	1624	TAGTTTATGTGCTCTCGTGAAGGCGTCTGGGTTTCAGCACCAAGAGGCTGGTGCCA 1683
DB	1896	TGCTCTATGTTTCTCGTGAAGGCGTCTGGGTTTCAGCATTCAGGATCACAAGAAAGCTGGTGCCA 1955
QY	1684	TGAATGCCCTGTCTCGTGTCTCAGCTGCTTACTAATGCAACAATACATGTTTCAATCTTG 1743
DB	1956	TGAATGCTCTGTCTCGTGTCTCAGCTGCTTACCAGTGGACATACATGTTGAATCTTG 2015
QY	1744	ATTGTGATCACTACATCAACAAACAGCAGGCTGTCCGAGAAGCTATGCTCTCTTAATGG 1803
DB	2016	ATTGTGATCACTACATCAACAAACAGTAAAGGCTCTCAGGGAAGCTATGCTCTCTTAATGG 2075
QY	1804	ATCCAAACCTAGGTCGCAAGTCTGTATGTGAGTTCCTCCACAAAGGTTTGTATGGGATGG 1863
DB	2076	ACCTTAACCTAGGAAAGGAGTGTCTGCTACGTCTCCAGGATTTCCAGAGATTCGATGGGATGG 2135
QY	1864	ATAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTTCATATATACTTGTGAGGGCC 1923
DB	2136	ACAGGATGATCGATATGCAACAGGAACACCGTGTGTTTTCGATATTAATCTTGAGAGGTC 2195
QY	1924	TTGAAGGCATTCAAGGACCAAGTTTATGTGGGAACCTGGTTGTGTTTTCACACAGAACAGCTA 1983

Query Match	64.9%;	Score 2352.4;	DB 7;	Length 3222;	
Best Local Similarity	83.8%;	Pred. No. 0;			
Matches 2675;	Conservative	0;	Mismatches 516;	Indels 3;	Gaps 1;
Qy	109	ACGTGTGCCAGATCTGCGCGAGCGGCGCTGGGACACACGTTGGAGCGGAGCGTCTTTCACCG	168		
Db	32	ATGCTGCCAGATCTGCGCGAGCGGCGTGGGACGCGCGGAGCTCTTTCACCG	91		
Qy	169	CCTGCGAGCTCTGCGGCTTCCCGGTCTGCGCGCCCTGCTAGCAGCAGCGGCGAGG	228		
Db	92	CCTGCGAGCTCTGCGGCTTCCCGGTCTGCGCGCCCTGCTAGCAGCAGCGGCGAGG	151		
Qy	229	GCACCCAGGCTGCTCCAGTGAAGCAAGCAAGTACAGCGCCACAGAGGAGCCCGCAGA	288		
Db	152	GCACCCAGGCTGCTCCAGTGAAGCAAGCAAGTACAGCGCCACAGAGGAGCCCGCAGA	211		
Qy	289	TCCCGGGGAGGAGGCGACGACACTGATGCCGATGATGGTAGTGACTTCAACTACCCCTG	348		
Db	212	TACTTGGGATGAAGCGATGATGTTGATCGGATGATGCTAGTGAATGAACTATCCAA	271		
Qy	349	CATCTGCACTGAGGACAGAGCAGAGATGCTGACAGATGCGCAGCTGGCGCATGA	408		
Db	272	CATCCGCAACACAGGACCAAGCACAAGATGCGGAGAGATGCTCACCTGGGCGCATGA	331		
Qy	409	ACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGCCCTCT	468		
Db	332	ACTCTGGGAGGATGATGATATGTTCCATTTCTAGTATGATGTTGAGATTGGTCACTC	391		
Qy	469	CCAAGTATGACAGTGGAGAGATCCCTAGGGGATAGCTCCCTTCAGTCAACCAACGCCAGA	528		
Db	392	CCAAGTATGACAGTGGTGAATCCCTCGCATATATATCCCATCTCACTCACAGCCAGA	451		
Qy	529	TGTCAGAGAAATCCCTGGAGCTTCGCTGATCATCATGATGTCCTCCCTACGGGGAACA	588		
Db	452	TCTCAGTGAATTCCTGGAGCGTCCCTGATCATATGATGTCCTCCGTTGGGAACATTG	511		
Qy	589	TCACAGACGTGCTCCCTTTCCCTATGTGAATCATTCACCAAAATCCGCTCAAGGAGTTCT	648		
Db	512	GCAGACGTGGCATCCATTTCCCTATGTGAACATTCACCAACCCATCAAGGAGTTT	571		
Qy	649	CGGCGATATTGGAAATGTTGCCCTGGAAAGAGAGATGATGGCTGGAAATGAAGCAGG	708		
Db	572	CTGTAGCTTGGCAATGTTGCATGGAAAGAGAGATAGTGGCTGGAAATGA--AGG	628		
Qy	709	ACAGGCTGCGATTCCTCATGATTAATGGGACAGCAAGTGTCCCTCTGAGGCTGGCGAG	768		
Db	629	ATAAGGTGCAATTCCTCATGCTAATGCTACTAGCATGTCTCTTCAGAAAGGCTGGGG	688		
Qy	769	CTACTGACATCGATGCTACTCTGAATACAAATGGAAGAGCGCTTTTACTGAATGATGAA	828		
Db	689	TTGCTGACATTTGATGCTACTCTGATATACATGAGGAGCGCTTACTGAATGATGAA	748		
Qy	829	CTGCGCAGCTCTATCTAGAAAAGTCCCATTTGCTTCCCTCAAAATAAATCCCTACAGAA	888		
Db	749	CACGCCAACCTCTATCGAGAAAAGTGCCATTTGCTATCCAGAAATAAATCCGTCACAGA	808		
Qy	889	TGCTCATTTGCTCGGTTGGTTGCTCTTAAGCATCTTCCGCTACGCTCTCAAAATC	948		
Db	809	TGGTATTTGCTCCGATTTGATTTGCTTATGATTAATCTTGCACCTACCGTATCAAAATC	868		
Qy	949	CTGTGCGTAATGATACCCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTGCTT	1008		
Db	869	CTGTGCGTAATGATACCCGCTATGTTGCTCTCTGTCATATGTGAGATTTGGTTGCTC	928		
Qy	1009	TATCTGTGATACATGATTCGGAAGTGGTTTCCAAATCAACCGGGAGACCTACCTTG	1068		
Db	929	TATCTGTGATTCGGAATTCGGAATTCGGAAGTGGTTTCCAAATCAACCGGAACTTACCTGG	988		
Qy	1069	ATAGACTGGCTTTAAGGTATGACGAGAAAGTGAACCGTCTCAGTTGGCTGCTGTGACA	1128		
Db	989	ATAGCTGGCTTTAAGGTATGACGAGAAAGTGAACCATCTCAATTTGGCTCCTGTGACA	1048		
Qy	1129	TATTTGTCAGTACAGTCGACCCCTTGAAGGAGCCACCTATCTGCTCACTGCCAACACTGTGC	1188		
Qy	1049	TTTTTGTGCACTGTGATCCTATGAAGAACTCTCTTGTCTCACTGCCAATCTGTGC	1108		
Qy	1189	TATCCATTTCTGTTGATTAATCCCGTGGACAGGCTCTCTGCTATGATATCTGATCAGC	1248		
Db	1109	TCTCCATCTCTGTTGGATTACCTGTTGATGAAGGTATCTGCTATGTTGCTGATGATG	1168		
Qy	1249	GAGCTCAATGCTGACTTTTGGACGCTTGGCTGAGACTTCAGAGTTTGTCTAGGAAATGGG	1308		
Db	1169	GAGCTCAATGTTTAAACATTTGATGCGCTTGTGAGACTTCAGAAATTTGCTAGAAATGGG	1228		
Qy	1309	TACCATTTGTGAAGAGATGATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTCCAGA	1368		
Db	1229	TACGTTCTCTAAAAAGTACAGCATAGAACCCAGAGCTCCAGAGTGTGTTTTCACAGA	1288		
Qy	1369	AAATTGATTTACCTGAAAGACAAAGTCCAGCTTCAATTTGTTTAAAGACCGCGGCGCATGA	1428		
Db	1289	AAATTGATTTACTTTGAAAGACAAAGTTCAGGCTTCTTTTGTAAAGATCGTCTGTCGCAATGA	1348		
Qy	1429	AGAGAGAAATGAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAAGCATTTGAAAG	1488		
Db	1349	AGAGGATATGAAGAAATTTAAAGTTCGTTTAAATGCTCTCTCGCCCAAGGCACAGAAG	1408		
Qy	1489	TCCCGAGGAGGATGGATCATGCAAGATGGCACACCATGCGCCAGGAAAACAATACAGGG	1548		
Db	1409	TTCTGTAGGAGGATGGATTAATGCAAGATGGCACACCTTGGCCCTGGTAAACAATACCAGAG	1468		
Qy	1549	ATCATCTCGAATGATTCAGGTTTCTCTGGTTCACAGTGGTGGCTTGTACTGTAGGTA	1608		
Db	1469	ACCATCTCGAATGATTCAGGTTTCTCTGGTTCATAGTGGAGGCTTGTATCCGAAAGCA	1528		
Qy	1609	ATGAGCTCCCGCTTATGTTTATGTGCTCTGTAAGACGCTCTGCTGGGTTCCAGCACCA	1668		
Db	1529	ATGAGCTTCTCGTTTGTAGTCTATGTGCTCTGTGAGAAAGTCTCTGGATTTCCAACACATA	1588		
Qy	1669	AGAAGGCTGTGCCATGAATGCCCTTGTTCGTGCTCAGTGTCCTTACTTAATGGAAT	1728		
Db	1589	AAAAGGCTGTGCTATGAATGCACTTCTGCTGTATCAGCTGTCTTACTTAATGGAAT	1648		
Qy	1729	ACATTTGATCTTGTATGTCATCATCAACACAGCAGCAGGCTGTCCGAGAACTA	1788		
Db	1649	ACTTGTGAAATCTTGTACTGTGATCATCAACAAATGCAAGCTCTCCGAGAGGCTA	1708		
Qy	1789	TGTGCTTCTTAAATGGATCCAAACCTAGGTCGCAAGTCTGTATGTGCAAGTCTCCACAAA	1848		
Db	1709	TGTGCTTCTTAAATGGATCCAAACCTAGGAGGCGTGTCTGTATGTCCAAATCTCCCTCAGA	1768		
Qy	1849	GGTTTGTAGGATTTGATAGGAATGATCGATATGCAAAACAGGAAACACTGTCTTTTGTGATA	1908		
Db	1769	GGTTTGTAGGATTTGATAGGAATGATCGATATGCAAAACAGGAAACACTGTCTTTTGTGATA	1828		
Qy	1909	TTAACTTGTAGGCGCTTGAAGGCAATTCAGGACAGTTTATGTGGGAACTGGTTGTGTTT	1968		
Db	1829	TTAACTTGTAGGCGCTTGAAGGCTTGAAGGCAATTCAGGACAGTTTATGTGGGAACTGGTTGTGTT	1888		
Qy	1969	TCAACAGAAACAGCTTATCTATGTTTATGAGCCCTCAATTAAGGCGGAGAAAGCAGGTTTCT	2028		
Db	1889	TCAACAGAAACAGCTTATCTATGTTTATGAGCCCTCAATTAAGGCGGAGAAAGCAGGTTTCT	1948		
Qy	2029	TGGCATCATATGTGGGGGCAAGAGGAGGCAAGCAAGTCAAGAAAGAGGCTCAGATA	2088		
Db	1949	TCTCTTCGCTTTTGTGGGGGCAAGAGGAGGCAAGAAAGTCAAGAGGAGGAGCAGGAAA	2008		
Qy	2089	AGAAAAAGTCAACAAAGTGTGACAGTCTGTGTTCCAGTATTTCAATCTCGAAGACATAG	2148		
Db	2009	AGAAAAAGTCAACAAAGTGTGACAGTCTGTGTTCCAGTATTTTAACTTTGAAATATAG	2068		
Qy	2149	AGGAGGCTGTGAAGGCTGCTGGGTTTGTATGATGAGAAATCAGTTCTCATGTCTCAATGA	2208		
Db	2069	AGGAGGCTGTGAAGGCTGCTGGGTTTGTATGATGAGAAATCAGTTCTCATGTCTCAATGA	2128		
Qy	2209	GCTTAGAGAGAGATTTGGCCAGTCAGCAGATTTTGTGCTTCCACTCTGATGGAATATG	2268		

Db 2129 GCTTAGAGAAAAGATTGGTCAATCTAGTGTGTTTTTGTAGCCCTCCACTCTGATGGAATATG 2188
QY 2269 GTGGTGTCTCAGTCTCCTCAGATCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCATPAA 2328
Db 2189 GTGGTGTCTCCTCAATCCGCAACTCCAGAAATCTCTTCTGAAGAAGGCCATACATGTTATCA 2248
QY 2329 GTTGTGGCTATGAGGACAAGTCTGAATGGGAACATGAGATTGGTTGGATCTATGGATCTG 2388
Db 2249 GCTGTGGCTATGAGACAAAGTGAAGTGGGAACATGAGATTGGATGGAATCTATGGTCTG 2308
QY 2389 TCACAGAAGATATCTTACTGGATTCAAGATGACGACGAGCTGGCGTTTCAGTCTATT 2448
Db 2309 TTACAGAAGATATCTTACTGGATTCAAGATGATGATGATGATGATGATGATGATGATGATG 2368
QY 2449 GCATGCCAAGCGCCAGCTTTCAAGGATCTGCCCCCAATCAATCTTTCAGATCTGCTGA 2508
Db 2369 GCATGCCAAGCGCCAGCAATTTCAAGGGGTCTGCCCAATTAATCTTTCGGATCTGCTGA 2428
QY 2509 ACCAAGTCTGCGGTGGGTCTCGGTTCGTGTTGAATCTCTTTTCAGCGGCAATTCGCCCT 2568
Db 2429 ACCAGGTGCTTCGGTGGGTCTCTGGTTCGTGGAATCTCTTTTCAGCGGCAATTCGCCCA 2488
QY 2569 TATGTTATGGCTACGAGGCGCCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCA 2628
Db 2489 TATGTTATGGGTACGAGGAGCGCTTAAGTTCCTGGAGAGATTCGCTTACATCAACACCA 2548
QY 2629 CCATTTACCACTAACCTCTCCCGCTCTAGTCTATTTGATATATTTGCTGCTATCTGTC 2688
Db 2549 CTAATTTACCACTACATGATTCACCTCTCTTGTACTGTCATATTTGCCAGCTATCTGTC 2608
QY 2689 TGCTCACTGGAAGTTCATGACGACAGATAGCAACTTGGCCAGTATCTGGTTCATG 2748
Db 2609 TTCTCACTGGAGGTTTCATACACAGAGATAGCAACTTGGAGATATTTGGTATTATCT 2668
QY 2749 CGCTCTCTCTTCAATTTTCGCACTGGTATCTCTGAGATGAGGTGGAGTGGTGGCA 2808
Db 2669 CTCTCTCTCTGTCATTTTGTACTGTTATCTTGGATGAGGTGGAGTGGTGGCA 2728
QY 2809 TTGAGAGTGGTGGAGGAATGAACAGTTCCTGGTCAATGGAGGTATCTCTGACATCTGT 2868
Db 2729 TCGATGAATGGTGGAGGAATGACAGTTCCTGGTCAATGGAGGTATTTCCGCTCATCTAT 2788
QY 2869 TTGCGGTCTTTCAGGCTCTCTGAGGTCTTGGCGGTATCGACACCACTTCATCTGCA 2928
Db 2789 TTGCGGTCTTCAGGCTCTCTGAGGTCTTGGAAATGATACCACTTCATCTGCA 2848
QY 2929 CTTCAAGCTTAATGACGAAGAGCGACTTTGCTGAGCTTACATGTTCAAGTGGACGA 2988
Db 2849 CTTCAAGGCTCTGATGAAGAAGCGATTTGCTGAGCTTACATGTTCAAGTGGACAA 2908
QY 2989 CGCTTCTATCCTCGGACGACATTTGATCATTTAAAGTGTGGTGTGCTGCTGGCA 3048
Db 2909 CCGTCTGATCCCGGACCACTATTTTATCATCAACCTGGTGGGTGTGTTGCTGGTA 2968
QY 3049 CTTCTACGCCATCAACAGTGTGTACCAATCATGGGGCGCTCTTTTGGAAAGCTCTTCT 3108
Db 2969 TCTCTACGCTATCAACAGTGGCTACCAATGATGGGTCCGCTCTTTGGAAAGCTCTTCT 3028
QY 3109 TTGCGCTTCTGGGTGATTTGCTACTATACCCATTTCTCAAGGCTCTTATGGGCGAGCAAA 3168
Db 3029 TTGCGCTTCTGGGTGATTTGCTACTATACCCATTTCTCAAGGCTCTCATGGGTGGCGACA 3088
QY 3169 ACCGACACCGAGGATTTGATGCTGGGTGCTGCTCTCTGCTGCTTCTATCTCTCTTCTG 3228
Db 3089 ACCGTACACCGACCAATTTGTTGTTTGGGCCATCTCTCTGATCGATCTTCTCTCTGCG 3148
QY 3229 TGTGGGTCTGTTGATTCATTCATTCACCTGCTGCTGCGCCCAATATCCAAACTGTCG 3288
Db 3149 TGTGGGTCTGTTGATTCATTCATTCACCCCGTGTACAGGCGCTGTACACCCCAAAATGTCG 3208
QY 3289 GCATCAACTGCTAG 3302
Db 3209 GTATCAACTGCTAG 3222

RESULT 7
AAZ99533
ID AAZ99533 standard; DNA; 3704 BP.
XX AC AAZ99533;
XX DT 03-JUL-2000 (first entry)
XX DE DNA encoding a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX OS Zea mays.
FH Key Location/Qualifiers
FT CDS 272..3496
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1544..1546, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
PN WO200009706-A2.
XX 24-FEB-2000.
XX 16-AUG-1999; 99WO-US018760.
XX 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI: 2000-224343/19.
XX P-PSDB; AAY84121.
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
PS Claim 1, Page 191-196; 119pp; English.
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting or
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of the present invention can also be used for recombinant expression of
XX their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein
SQ Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 0 U; 1 Other;

Query Match 64.1%; Score 2325.2; DB 3; Length 3704;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 544; Indels 9; Gaps 3;

QY 66 GCGAGCGGACGCGCTTGAAGTTCGGGAGGACCGGGCCGGGAGGTGTGCCAGATCTGC 125
Db |||||
QY 275 GACGGCGCGACCGCACGAATTCGGGGAAGCATGTGGCCGGGCGAGGTGTGCCAGATCTGC 334
Db |||||
QY 126 GCGAGCGGCTTGGGACACAGCTTGGAGCGGAGCGTCTTACCGGCTGCGACGTGTGCCG 185
Db |||||
QY 335 GCGAGCGGCTTGGGACACCGCGGGGACGGCGACCTCTTACCGGCTGCGACGTGTGCCG 394
Db |||||
QY 186 TTCCGGTCTGCGGCCCTGTACGACGACGAGCGGAAGAGGAGGACCCAGGCTGCTC 245
Db |||||
QY 395 TTCCCGTGTGCGGCCCATGCTACGAGTACGAGCGCAAGGACGCCACCCAGGCGTCCCG 454
Db |||||
QY 246 CAGTGAAGACCAAGTAAAGCGCCACAGAGGAGCCAGCGATCCCGCGGGAGGAAGGC 305
Db |||||
QY 455 CAGTGCMAAGACTAAGTACAAGCGCCACAAGGAGGCCACAGTACACGGTGAAGAAAT 514
Db |||||
QY 306 GAGGACACTGATGCCGATGTAGTGTGACTTCAACTACCTGATCTGCACTGAGGAC 365
Db |||||
QY 515 GAGGATGTGATGTGACGATGTGAGTGACTTACAACCTACCAAGCATCTGGCAACAGGAT 574
Db |||||
QY 366 CAGAGCAGAGATTGCTGACAGATGGGAGCTGGGCGCATGACACCGGGGCGAGTGGC 425
Db |||||
QY 575 CAGAGCAAAAGATTGTGAGAGATGCTCACTTGGCGGCAAACTCACGTGGCGAGTG-- 632
Db |||||
QY 426 AATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAAAGTATGACAGTGA 485
Db |||||
QY 633 -ATATTGGCTTGGCTAAGTATGACAGCGGTGAATTTGGGCATGGGAAGTATGACAGTGGT 691
Db |||||
QY 486 GAGATCCCTAGGGGATAGCTCCCTTCAGTCAACCAAGCGAGATGTGAGGAGAAATCCCT 545
Db |||||
QY 692 GAGATCCCTCGTGGATATATCCGTCACCTAACTCATAGCCAGATCTCAGGAGAGATTCCT 751
Db |||||
QY 546 GAGCTTCGCTGATCATCATGATGTCCCTACGGGGAACATCAGCAGACGTGCTCCG 605
Db |||||
QY 752 GAGCTTCCTCGTATCATATGATGTCTCTTGGGAACATTTGGCAGCGGTGGACATCAA 811
Db |||||
QY 606 TTTCCTTATGTGAATCAATTCACCAAAATCCGTCAAGGAGTGTCTCCGCGAGTATTTGGGAAT 665
Db |||||
QY 812 TTTCCTTATGAATCATCTCCAAACCCATCGAGGAGTGTCTCCGCTGACCTTGGCAAT 871
Db |||||
QY 666 GTTGCTCGAAAGAGAGATGTATGTGTGAAATGAAGCAGGACAAAGGTTGCAATCCC 725
Db |||||
QY 872 GTTGCAATGAAGAGAGGTTGGATGGATGGAATGA--AGGATAAAGGTGCAATTCCT 928
Db |||||
QY 726 ATGACTAATGGCAAGCATGTCTCCCTCAAGGTGGGCGAGTACTGACATCATGATGCA 785
Db |||||
QY 929 ATGACCAATGGAAACAGCATTTGCTCCATCAGAGGGCGTGGAGTTGCTGATATTGATGCT 988
Db |||||
QY 786 TCTACTGAATACAAACATGGAAGCGCTTTACTGAATGATGAAACTCGCCAGCCTCTATCT 845
Db |||||
QY 989 TCTACTGATTAACATGGAAGATGCTTACTGAATGATGAAACTCGGCAACCTCTATCT 1048
Db |||||
QY 846 AGAAAGTCCCAATGTCTCTCCAAATAAATCCCTACAGATGGTCAATGTTCTGCGG 905
Db |||||
QY 1049 AGAAAGTGGCAATTCCTTATCCAGAAATAAATCCGTACAGATGGTCAATTTGCTACGT 1108
Db |||||
QY 906 TTGGTGTCTTAAGCATTTCTGCACTACGCTCTCAAAATCCGTGCGGTGAATGCAATAC 965
Db |||||
QY 1109 TTGGCTGTCTATGCAATTTCTTGGCTACCGTATCACATCTGTGAACAATGCAAT 1168
Db |||||
QY 966 CCACTGTGGCTTTATCTGTTATATGTGAGATTTGGTTTGGTTTATCTCGATACTGGAT 1025
Db |||||
QY 1169 CCACTGTGGCTTTATCCGTCAATGTGAGATCTGGTTTGGTTTCTTGGATTTTGGAT 1228
Db |||||
QY 1026 CAGTTCGGAAGTGGTTTCCAAATCAACCGGAGACCTACCTTGTATAGACTGGCTTTAAG 1085
Db |||||
QY 1229 CAGTTCGCAAGTGGTCCCAATCAACCGTGAACATACATCTTGTATAGACTGGCTTTAAG 1288
Db |||||
QY 1086 TATGACGAGAGTGAACGCTCTCAGTTGGCTCTGTGACATATTTGTGACGACAGTC 1145
Db |||||
QY 1289 TATGACCGAAGAGTGAACCATCTCAATTAGTCTCTGTGATATTTTGTGACGACTGTG 1348
Db |||||

QY 1146 GACCCCTTGAAGGAGCCACCTATCGTCACTGCCAACACTGTGCTATCTCATTTCTGTCTGT 1205
Db |||||
QY 1349 GATCCAATGAAGGAGCTCTCTGTGACACTGCAAACTACTGTGCTTTCCATCTTGTCTGTC 1408
Db |||||
QY 1206 GATTATCCCGTGGACAGAGTCTCTTGTCTATGTATCTGTATGATGACGAGCTTCAATGCTGACT 1265
Db |||||
QY 1409 GATTATCCGGTTGACAGGATCTTGTCTATGTTTCCGATGATGAGCTGTATGCTGACT 1468
Db |||||
QY 1266 TTTGAGCGATTCGCTGAGACTTCAGAGTTTGTAGGAAATGGGTACCATTGTTGGAAGAAG 1325
Db |||||
QY 1469 TTTGATGCTCTCTGAACTTCAGAGTTTGTGTAAGAAATGGGTCCGTTCTGTGAAGAAG 1528
Db |||||
QY 1326 TATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTTCOCAGAAAATGATTAACCTGAAA 1385
Db |||||
QY 1529 TACAACATAGAGCTTANGGCCCGGAATGCTATTGCTCAGAAAATGATTAACCTGAAA 1588
Db |||||
QY 1386 GACAAAGTCAGCTTCAITTTGTTAAAGCCGCGGCCCATGAAGAGAAATATGAGAA 1445
Db |||||
QY 1589 GACAAAGTTCAAACCTCATTTGTAAAGAACGCGCGGCCCATGAAGAGAAATATGAGAA 1648
Db |||||
QY 1446 TTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAAGTCCCGAGGAGGATGG 1505
Db |||||
QY 1649 TTTCAAGTTCGATCAATGCTCTTTAGCCAGGCAAAAAGTTCCCGAGGAGGATGG 1708
Db |||||
QY 1506 ATCATGCAAGATGGCAGCACCATGCGCAGGAAACAAATACAGGAGTCACTCTGGAATGAT 1565
Db |||||
QY 1709 ATCATGCAAGATGGTACACCTTGGCCTGGGAACAATACTAGGACCATCTCGGAATGAT 1768
Db |||||
QY 1566 CAGGTTTCTTGGTCAAGTGTGGCTTGTACTGTAGGGTAAATGAGTCCCGCTTTA 1625
Db |||||
QY 1769 CAGGTTTCTTGGTCAAGTGTGGGCTTGAAGGCAATGAATTCCTCGTTG 1828
Db |||||
QY 1626 GTTTATGTGCTCGTGAAGAGCGTCTGGGTTCCAGCACCAAGAGGCTGGTGGCCATG 1685
Db |||||
QY 1829 GTTTATGTGCTCGTGAAGAGCGTCTGGATTTCAACATCACAAGAGGCTGGTGGCCATG 1888
Db |||||
QY 1686 AATGCCCTTGTTCGTGCTCAGCTGCTCTTACTAATGGAACAATACATGTTGAAATCTTGAT 1745
Db |||||
QY 1889 AATGCATTTGTTGTTGATCAGCTGCTCTTACTAATGGCAATACATGTTGAATCTTGAT 1948
Db |||||
QY 1746 TGTGATCACTACATCAACACAGCAGGCTGTCCGAGAGCTATGCTTCCCTAATGAT 1805
Db |||||
QY 1949 TGTGACCACTACATCAATATATGCAAGGCTCTTCGAGAGCTATGCTTCTTATGGAC 2008
Db |||||
QY 1806 CCAACCTAGTCCGCAAGTCTGTATGTGAGTCCCAAAAGGTTTGAATGGATGAT 1865
Db |||||
QY 2009 CCAACCTAGAGAGATGTCTGTATGTCCTTCTCAGAGGTTTGAATGGATGAT 2068
Db |||||
QY 1866 AGAATGATCGATATGCAAAACAGAACACTGTCTTTTTTGAATATACTTGAAGGCGCTT 1925
Db |||||
QY 2069 AGAATGACCGATATGCAAAACAGAACACTGTGTGTTTTTCGATATTAACCTTGAGAGTCTT 2128
Db |||||
QY 1926 GACGGCAATTCAGGACCGAGTTTATGTGGAACTGTTGTTTTTCAACAGAACAGCTATC 1985
Db |||||
QY 2129 GACGGCAATTCAGGCGCAGTTTATGTGGAACTGTTGTTTTTAAACAGAACCGCTTAA 2188
Db |||||
QY 1986 TATGTTTATGAGCCCCCAATTAAGCGGCAAGAACGAGGTTTCTTGGCATCACTATGTGGG 2045
Db |||||
QY 2189 TATGTTTATGAGCTTCCAGTCAAGAAAAAAGAGCCAGGCTTCTTCTCTCGCTTGTGGG 2248
Db |||||
QY 2046 GCGAAGAGAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAAAGTCAAGCAAG 2105
Db |||||
QY 2249 GGAAGGAAAAAGACGTCAAAATCTAAGAA---GAGCTCGGAAAAAGAAAGTCAATAGA 2305
Db |||||
QY 2106 CATGTGACAGTCTGTGTTCCAGTATTTCAATCTCGAGCATAGAGGAGGTTGTTGAAGGT 2165
Db |||||
QY 2306 CAGCAGACAGTCTGTGACCAATTTAATCTCGAGATATAGAGGAGGATTTGAAGGT 2365
Db |||||
QY 2166 GCTGGGTTTGTATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTTTAGAGAGATTT 2225
Db |||||
QY 2366 TCTCAGTTTGTATGAGAAATCGCTGATTTATGTCTCAATGAGCTTTGGAGAGAGATTT 2425
Db |||||
QY 2226 GGCAGTCAAGCAGATTTGTTGCTCCACTCTCTGATGGAATATGTTGGTGTCTCTCAGTCC 2285
Db |||||

||||| GGCAGTCCAGTGTGTTTGTAGCCTCTACTCTGATGGAATATGTTGGTTCCTCCAAATCT 2485
2486 TCCACTCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCAATAAGTTGTGGCTATGAGAAC 2345
2486 GCAACTCCAGAGTCTCTTCTGAAGAAGCTATTCATGTCTCATGCTGTGGCTATGAGGAC 2545
2346 AAGTCTGAATGGGAACTGAGATTTGGTGGATCTATGATCTGTACAGAGATATTTCT 2405
2546 AAAACTGACTGGGAACAGATTTGGTGGATCTATGTTCTGTATACAGAGACATTTCTC 2605
2406 ACTGATTCAGATGACGCAAGAGGCTGGCTTCAGTCTATTTGATGCCAAGCGCCCA 2465
2606 ACCGATTCAGATGATCTGCTGAGGCTGGCAATCACTCTGATGCTTAAGGACCA 2665
2466 GCTTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTGGGTGG 2525
2666 GCTTTCAAGGATCTGCTCTTATCAACTTTTCGGATCGTTTGAATCAAGTGTCTGGTGG 2725
2526 GCTCTGGTCTGTGAAATCTTTTCAGCGGCAATGCCCCCTTATGTTATGCTTACCGA 2585
2726 GCTCTGTGTTCAATGAAATCTTTTCAGAGGCAATGTCATATCAACAAATTTATCCATCA 2785
2586 GGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCACTTTTACCCTAAACC 2645
2786 GCGCGCTTAAATCTCTGGAGAGATTTGCTTATATCAACAAATTTATCCATCA 2845
2646 TCTCTCCGCTCTAGTCTATTTGATATGCTGCTATCTGCTCTCACTGGAAGTTC 2705
2846 TCAATCCCGCTCTCTCTGCTACTGCTATTTGCGAGAGTTTGTCTTCACTGGGAAGTTC 2905
2706 ATCATGCCAGATTTAGCAACTTGGCCAGTATCTGGTTCATTCGCTCTTCTTCAATT 2765
2906 ATCATCCCAAGATTTAGTAACCTAGAGAGTGTGTTGTTTATATGCTCTTTATCAATC 2965
2766 TTGCGCACTGTPATCTCTGAGATGAGTGGAGTGTGTTGCAATTCAGAGTGTGGAGG 2825
2966 TTGCGCACTGTPATCTCTGAGATGAGTGGAGTGTGTTGCAATTCAGAGTGTGGAGG 3025
2826 AATGAACAGTCTGGGTCTATGAGATATCTCTGACATCTGTTGCGCTCTTTCAGGCT 2885
3026 AACGAGAGTCTGGGTCTATGAGTGTGTTGCAATTCAGAGTGTGGAGG 3085
2886 TTTGAGAGTCTGGGTCTATGAGATGAGTGGAGTGTGTTGCAATTCAGAGTGTGGAGG 2945
3086 TTTGAGAGTCTGGGTCTATGAGATGAGTGGAGTGTGTTGCAATTCAGAGTGTGGAGG 3145
2946 GAAGAAGGCTCTGCTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCCTCGG 3005
3146 GAAGAAGTGTATTTGCGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCCTCGG 3205
3006 ACGACATTTTATCAATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3065
3206 ACCACTATTTTATCAATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3265
3066 AGTGTGTTACCAATCATGAGGCGCTCTTTTGGAGCTCTTTTGGAGCTCTTTTGGAGCT 3125
3266 AGCGGTACCAATCATGAGGCGCTCTTTTGGAGCTCTTTTGGAGCTCTTTTGGAGCT 3325
3126 GTTCACTTATACCATCTCTCAAGGCTCTTATGAGGAGGCAAAACCGCACCGAGCAT 3185
3326 GTTCACTTATACCATCTCTCAAGGCTCTTATGAGGAGGCAAAACCGCACCGAGCAT 3385
3186 GTTCACTTATACCATCTCTCAAGGCTCTTATGAGGAGGCAAAACCGCACCGAGCAT 3245
3386 GTTCACTTATACCATCTCTCAAGGCTCTTATGAGGAGGCAAAACCGCACCGAGCAT 3445
3246 CATTCACCTACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3304
3446 CATTCACCTACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3504

AAZ58271
ID AAZ58271 standard; cDNA; 1734 BP.
XX
AC AAZ58271;
XX
DT 08-MAY-2000 (first entry)
XX
DE Wheat cellulose synthase cDNA contig.
XX
KW Wheat; cellulose synthase; transgenic plant; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 2..1522
FT /*tag= a
FT /partial
XX
PN WO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015871.
XX
PR 14-JUL-1998; 98US-0092844P.
XX
PA (DUFO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
DR WPI: 2000-182431/16.
DR P-PSDB; AAY58840.
XX
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
probes for isolating cDNAs and genes encoding homologous proteins, for
producing transgenic plants.
XX
PS Claim 3; Page 66-67; 93pp; English.
XX
CC The present sequence is that of a contig of cDNA clones and encodes a
portion (see AAY58840) of wheat cellulose synthase (CS). The cDNA clones
were isolated from wheat 7-day seedling leaf and 7-day etiolated seedling
root tissue cDNA libraries on the basis of homology to Arabidopsis and
cotton CS sequences. The invention relates to isolated nucleic acid
fragments encoding plant CS and to CS polypeptides. It also relates to
the construction of a chimeric gene encoding all or a portion of the CS,
in sense or antisense orientation, where expression of the gene results
in altered levels of the CS in transformed host cells. The host cells can
be used to screen compounds for their ability to inhibit CS activity. CS
nucleic acids are also useful for producing transgenic plants having
altered levels of CS, and hence altered levels of fibre. CS may also
serve as a target for the development of novel herbicides
XX
SQ Sequence 1734 BP; 441 A; 369 C; 420 G; 502 T; 0 U; 2 Other;
Query Match 46.4%; Score 1691.6; DB 3; Length 1734;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1782 GAAGCTATGCTGCTTCTCTAATGGATCCAAACCTAGGTCCGCAAGTCTGTTATGTGAGTTC 1841
Db 2 GAAGCTATGCTGCTTCTCTAATGGATCCAAACCTAGGTCCGCAAGTCTGTTATGTGAGTTC 61
QY 1842 CCACAAAGTTTGTATGGGATTTGATAGGAATGATCGATATGCAACAGGAACTGTCTTT 1901
Db 62 CCACAAAGTTTGTATGGGATTTGATAGGAATGATCGATATGCAACAGGAACTGTCTTT 121
QY 1902 TTTGATATTAATCTGAGGGCCCTTCAGCGCATTCAGGCACTTATGTGGAACTGCT 1961
Db 122 TTTGATATTAATCTGAGGGCCCTTCAGCGCATTCAGGCACTTATGTGGAACTGCT 181
QY 1962 TGTGTTTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAAGCGGAAGGCCA 2021

Db 182 TGTGTTTTCAAGAACGGCTAICTATGTTATGAGCCCCCAATTAAGGCGAAGGCCA 241
Qy 2022 GGTTCCTTGTCATCATATGTTGGGGGCAAGAGCGCAAGTCAAGAAAGAGGC 2081
Db 242 GGTTCCTTGTCATCATATGTTNGGGGCAAGAGCGCAAGTCAAGAAAGAGGC 301
Qy 2082 TCAGATAAGAAAAGTCAAGCATGTGACAGATTCCTGTCAGATTAATCAATCTCGAA 2141
Db 302 TCAGATAAGAAAAGTCAAGCATGTGACAGATTCCTGTCAGATTAATCAATCTCGAA 361
Qy 2142 GACATAGAGGAGGTGTGTAAGTGTCTGGGTGTTGATGATGAGAAATCAGTTCTCATGTCT 2201
Db 362 GACATAGAGGAGGTGTGTAAGTGTCTGGGTGTTGATGATGAGAAATCAGTTCTCATGTCT 421
Qy 2202 CAATAGAGCTTAGAAGAGATTTGGCCAGTCAGCAGCATTTGTCCTCCACTCTGATG 2261
Db 422 CAATAGAGCTTAGAAGAGATTTGGCCAGTCAGCAGCATTTGTCCTCCACTCTGATG 481
Qy 2262 GAATATGTTGTTCTCAGTCTCCACTCCAGAACTCTCTTTTGAAGAAGCTATCCAT 2321
Db 482 GAATATGTTGTTCTCAGTCTCCACTCCAGAACTCTCTTTTGAAGAAGCTATCCAT 541
Qy 2322 GTCAATAGTTGTGCTTAGAGGACAAGTCTGAATGGGGAACAGATTTGTTGATCTAT 2381
Db 542 GTCAATAGTTGTGCTTAGAGGACAAGTCTGAATGGGGAACAGATTTGTTGATCTAT 601
Qy 2382 GGATCTGTACAGAGATTTCTTACTGGATTCAAGATGACGCAAGAGCTGGCGTTCA 2441
Db 602 GGATCTGTACAGAGATTTCTTAACTGGATTCAAGATGACGCAAGAGCTGGCGTTCA 661
Qy 2442 GTCTATTGTCATGCCAAGCGCCAGCTTTCAAGGGATCTGCCCACTCAATCTTTCAGAT 2501
Db 662 ATCTATTGTCATGCCAAGCGCCAGCTTTCAAGGGATCTGCCCACTCAATCTTTCAGAT 721
Qy 2502 CGTCTGAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATTCCTTTTCAGCGGCAT 2561
Db 722 CGTCTGAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATTCCTTTTCAGCGGCAT 781
Qy 2562 TGCCCTTATGTTATGCTACGAGGCGGCTCAAGTTCTGAGAGATTCGCTTATCATC 2621
Db 782 TGCCCTTATGTTATGCTACGAGGCGGCTCAAGTTCTGAGAGATTCGCTTATCATC 841
Qy 2622 AACACCACTATTACCACTAACTCTCTCCCGCTTCTAGTCTATTGATATTTGCCGTCT 2681
Db 842 AACACCACTATTACCACTAACTCTCTCCCGCTTCTAGTCTATTGATATTTGCCGTCT 901
Qy 2682 ATCTGTCTGCTCACTGGAAGTTTCATGTCGCGAGATTAAGCACTTTGGCCAGTATCTGG 2741
Db 902 ATCTGTCTGCTCACTGGAAGTTTCATGTCGCGAGATTAAGCACTTTGGCCAGTATCTGG 961
Qy 2742 TTCAATTCGCTCTTCTTTCAATTTTTCGCACTGTTATCTCTGAGATGAGGTGAGTGT 2801
Db 962 TTCAATTCGCTCTTCTTTCAATTTTTCGCACTGTTATCTCTGAGATGAGGTGAGTGT 1021
Qy 2802 GTTGGCAATTGACAGTGTGGAGGAATGAACAGTCTTGGGTCAATTGAGGATATCTCTGCA 2861
Db 1022 GTTGGCAATTGACAGTGTGGAGGAATGAACAGTCTTGGGTCAATTGAGGATATCTCTGCC 1081
Qy 2862 CATCTGTTGCGCTCTTTCAGGCTCTTGAAGGTGTTGCCGATATGACACCAACTTC 2921
Db 1082 CATCTGTTGCGCTCTTTCAGGCTCTTTCAGGCTCTTTCAGGATATGACACCAACTTC 1141
Qy 2922 ACTGTCACTCAAGGCTAATGACGAAGAGCGACTTTGCTGAGCTCTACATGTTCAAG 2981
Db 1142 ACTGTCACTCAAGGCTAATGATGAAGAGGCGACTTTGCTGAGCTCTACATGTTCAAG 1201
Qy 2982 TGGACAGCTTCTCATCCCTCCGACGACCAATTTGATCATTAACATGTTGTTGCTGCT 3041
Db 1202 TGGACAGCTTCTCATCCCTCCGACGACCAATTTGATCATTAACATGTTGTTGCTGCT 1261
Qy 3042 GCTGGACCTCTTACGCCATCAACAGTGGTTACCAATCATGGGGCGCTCTTTTGGGAAG 3101

Db 1262 GCTGTTACCTCTACGCCATCAACAGTGGTTTACCAATCATGGGGCGCTCTTTTGGGAAG 1321
Qy 3102 CTCTTCTTTGCTTCTGGGTGATTTGTTCACTTATACCATTCTCAAGGGTCTTTATGGGC 3161
Db 1322 CTCTTCTTTGCTTCTGGGTGATTTGTTCACTTATACCATTCTCAAGGGTCTTTATGGGC 1381
Qy 3162 AGCAAAACCGCACACCGCAGATTGTCATCGTCTGGGCTGCTCTCTCTCTCTCTCTCT 3221
Db 1382 AGCAAAACCGCACACCGCAGATTGTCATCGTCTGGGCTGCTCTCTCTCTCTCTCTCT 1441
Qy 3222 TCCTTGTCTGGGTTCGTGTTGATCCATTCACTACCGCTCTCGTGGCCCAAAATATCCAA 3281
Db 1442 TCCTTGTCTGGGTTCGTGTTGATCCATTCACTACCGCTCTCGTGGCCCAAAATATCCAA 1501
Qy 3282 ACCTTGGCATCACTCTAGGAAGTGGAGTCTTGTAGACAGAGAAAATATAACAGTGA 3341
Db 1502 ACCTTGGCATCACTCTAGGAAGTGGAGTCTTGTAGACAGAGAAAATATAACAGTGA 1561
Qy 3342 TCGAGCAACAACCCCGGAGCCAGAGAAATATTTATGTTGGGTTGTGAATTACTACGTTT 3401
Db 1562 TCGAGCCACCACTGTGGAGCCAGAGAAATATTTATGTTGGGTTGTGAATTACTACGTTT 1621
Qy 3402 GAGAAAGTTGTCAAAATTTGAGAAAACAACATTTGTAATAGATGTAATPAGACCATCTACCG 3461
Db 1622 GAGAAAGTTGTCAAAATTTGAGAAAACAACATTTGTAATAGATGTAATPAGACTATCTACCG 1681
Qy 3462 TTTTTCATGAGTTAAGCTCTTCTTTTGTGGAACAAGGAA 3502
Db 1682 TTTTTCATGAGTTAAGCTCTTCTTTTGTGGAAAAAA 1722

RESULT 9
AAC49550
ID AAC49550 standard; DNA, 3198 BP.
XX AAC49550;
XX AC AC
XX XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134769P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.

Query Match 46.3%; Score 1677.6; DB 3; Length 3198;
Best Local Similarity 72.0%; Pred. No. 0;
Matches 2330; Conservative 0; Mismatches 834; Indels 72; Gaps 8;

PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161921P.
PR 28-OCT-1999; 99US-0161931P.
PR 29-OCT-1999; 99US-0162142P.

77 CGCCCTGAAGTCGGGAGGACGCGGCGCGGACGCTGCGCAGATCTGCGCGACGCGCT 136
21 CGCGGAAAGCCGATGAAGAACATGTTCCGACACTTGCAGATCTGTAGTGAAATGT 80
137 GGGCACACAGTTGGACCGGACGCTCTTCAACGCTCGGACGCTGCGCGCTTCCCGGCTG 196
81 TGGCAAGACTGTTGATGGAGATCGTTTGTGGCTTGTGATATTTGTTCAITTCGCCAGTTG 140
197 CGGCCCTGTACGAGCACAGCGCAAGGAGGACCGACCGCTGCGCTCGAGTCAAGAC 256
141 TCGGCCCTGTGACGAGTATGAGAGGAAAGATGGGAATCAATCTTGTCTCAGTGCAGAAC 200
257 CAAGTCAAGCGCCACAGAGGAGCCGACGATCCGCGGAGGAGGAGGAGGACGACACTGA 316
201 CAGATCAAGAGGCTCAAGGTAGTCTGCTATTTCTGGTGATTAAGACGAGGATGGCTT 260
317 TGGCGATGATGGTA---GTGACTTCAACTACCTCGATCTGGCATCTGGCATGAGGACCAAGACA 373
261 AGCTGATGAAGTACTGTGAGTTCAACTACCT-----CAGAAGGA 302
374 GAAGATTGCTCAGAGGATGGCAGTGGCGCATGAACACCGGGGAGTGGCATGTGG 433
303 GAAATTTACAGAGCGGATGCTGGTGGCATCTTACTCGTGGGAGGAGGAAATGGG 362
434 CCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCAAGTATCACAGTGGAGATCCC 493
363 GBAACCCAGTATGATAAGAG-----TCTCTCAATCATCTTCTCTCTCTAC 413
494 TAGGGGATACGTCCTTTCAGTCAACAGCCAGATGTGAGGAAATCCCTGGAGCTTC 553
414 GAGCAGACAAG-----ATACTTCAGGAGAGTTTCTGCTGCTC 452
554 GCCTGATCATCATGATGTCCTTACCGGGAAATCAGCAGAGCTGCTCGGTTCCCT- 612
453 ACCTGAAGCGCTCTCTGATCTCTACTATGCTGGGGAAAGCGCTTCCCTATTATC 512
613 --ATGTGAATCATTCACCAATCCGTCAAGGAGTCTCCGGCAGTATGGGAATGTTC 670
513 AGATGTCAATCAATCAACCAATAGAAGATTGTGGATCTTGGACTCGGGAATGAGC 572
671 CTGGAAGAGAGAGTTGATGGCTGGAAATGAAGCAGCAAGGGTCCGATGCCATGAC 730
573 TTGGAAGGAGAGAGTTGATGGCTGGAAATGAAGCAAGAAATACTGGTCTCTGTCAG 632
731 TAATGGGACAAGCATTCCTCCTGTAAGGTCCGGGAGCTACTGACATCGATGATCTAC 790
633 CACGAG-----GCTGCTTCTGAAGAGGTGGAG---TAGATATTGATGCCAGC 680
791 TGAATCAACATGAAGACGCTTATCTGAATGATGAACACTCGCAGCGCTCTATCAGAAA 850
681 AGATATCCTAGCAGATGAGGCTCTGCTGAATGACGAAGCGAGGCGAGCTCTGTCAAGGAA 740
851 AGTCCCATGCTTCTCCAAATTAATCCCTACAGAAATGGTCAATTTGTTCTGCGGTTGGT 910

Db 741 AGTTTCAATTCCTTCATCAGGATCAATCCTTTACAGAAATGCTTATTATGCTGCGGCTGT 800
Qy 911 TGTTCTAAGCATCTTCTCTGCACTACCGTCTCAAAATCTGTGGTGAATGATPACCCACT 970
Db 801 TATCCTTTGTCTCTTCTTGTGATTAACCGTATAACAAACCCAGTGCCTTTGTCTCT 860
Qy 971 GTGGCTTTTATCTGTTATATGAGATTTGTTGTTGTTTCTTCTGGAATCTGGAATCAGTT 1030
Db 861 ATGGCTGGTCTCTGTGATATGAGATCTGTTTGGCTTATCTCTGGAATTTTGGATCAGTT 920
Qy 1031 CCCGAAGTGGTTTCCAAATCAACCGGAGACCTACTCTTGATAGACTGGCTTTAAGGATGA 1090
Db 921 TCCCAAGTGGTTTCTCTGTAACCGTGAAACCTACTCGACAGGCTTCTTTAAGATAGA 980
Qy 1091 CCGAAGGTGAACCGTCTCAGTTGGCTGCTGTGATGATATTTGTCTAGTACAGTGCACCC 1150
Db 981 TCGTGAAGGTGAGCCATCAGAGTTAGCTGTGTTGACATTTTCGTGAGTACTGTGTGCC 1040
1151 CTTGAAGGAGCCACTATCGTCACTGCCAACACTGTCTATCCATTCCTTGTGTGATTA 1210
1041 CTTGAAGGAGCCACCCCTTGTGACGCCAACACAGTCTCTCTATTTCTGGCTGTGACTA 1100
Qy 1211 TCCCGTGGACAAGTCTCTTCTGATGATCTGATGACGAGCTTCAATGCTGACTTTTGA 1270
Db 1101 CCCAGTTGACAAGGTGCTGTTATGTTCTGATGATGTTGCTGCTGCTATGTTATCATTTGA 1160
1271 CGCATTTGGCTGAGACTTTCAGAGTTTGTAGGAATGGTACCATTTTGTGAGAGTATGA 1330
1161 ATCATTTCGAGAAACATCAGAGTTTGTCTGTAATGGGTACCATTTTGCAGAAATATAG 1220
Qy 1331 CATTGAACCCAGAGCTCCCGAGTTTACTTTTGCAGAAAATGATTACCTGAAAGACAA 1390
Db 1221 CATAGAGCTCTGCGACAGAAATGGTACTTTTGTGCGAATAATAGATTACTTTGAAGATA 1280
Qy 1391 AGTCCAGCTTCATTTGTTTAAAGACCGCGGCGCATGAAGAGAAATATGAAGAAATTA 1450
Db 1281 AGTTCAGACATCATTTGTCAAAGATCGTAGAGCTATGAAGAGGAAATATGAGAAATTA 1340
1451 AATCAGATTAATGCGCTAGTTTCTAGGCATTTGAAAGTCCCGAGGAGGATGGAATCAT 1510
1341 AATCCGAATCAATGCACTTTGTTCCAAAGCGCTTAAATGTCTGGAAGAGGTTGGTTAT 1400
1511 GCAAGATGGCACACCATGGCCAGAAACAATACAGGGATCATCTCGAATGATTCAGGT 1570
Db 1401 GCAAGATGGCACACCGTGGCTGGAATATAACAAGGACCAATCCAGGAATGATCCAGT 1460
Qy 1571 TTTCTTTGGTCAAGTGGTGGCTTGATATGAGGGTAAATGAGCTCCCCGTTTGTGTTA 1630
Db 1461 CTTCTTAGGGCAAAATGGTGGACTTGATGACAGGGCAATGAGCTCCCGCTTTGGTATA 1520
1631 TGTGCTCTCGTGAAGAGCTCTGGTTCAGGACCAACAAGAGCTGTCCTGCAATGC 1690
1521 TGTTCTCGAGAAAGCGACAGGATTCAGACCCAAAGAGCTCGTGCTATGAATGC 1580
Qy 1691 CTTGTTCTCGTCTCAGCTGTCTTAAATGACAAATACATGTTGAATCTTGAATCTGTA 1750
Db 1581 ACTGGTGAAGTTCAGCAGTCTTACCAATGGACCTTTTCATCTTGAATCTTGAATGTA 1640
Qy 1751 TCATCATCAACAACAAGAGGCTGTCCGAGAGCTATGCTTCTTAATGGAATCCAAA 1810
Db 1641 TCATTCATTAATAACAGCAAGCCCTTAAGAGAGCAATGTGCTTCTGATGGACCCAAA 1700
1811 CCTAGGTCCCAAGTCTGTTATGTCAGTTCCCAAAAGTTTGTATGGGATTCGATAGAA 1870
Db 1701 CTTCCGGAAGCAAGTTGTTATGTTCTAGTTCCCAAAAGATTTGATGGTATCGATAGAA 1760
1871 TGATCGATATGCAAAACAGGAACACTGCTTTTTTTTGTATATTAACCTGAGGCGCTTGACGG 1930
Db 1761 CGATAGATATGCTAATCGTAATACCGTGTCTTTGATATTAACCTGAGAGGTTTAGATGG 1820
Qy 1931 CATTCAAGGACAGTTTATGTGGAACTGTTGTTGTTTCAACAGAACACGATCTATAGG 1990
Db 1821 GATTCAGGACCTGTATATGTCGGAATGTTGTTTCAACAGAACAGCATTTATACGG 1880

QY 2058 GCAAGCAAGTCAAAGAAAGAGCTCAGATAGAAAAGTCAAGCAAGCATGTGACAGT 2117
Db 2170 AAGAAATCCAAAGCTAAGAAAGAGTGGACAAAAGAAATCAGGAGGATACTGACTCA 2229
QY 2118 TCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGGTGTGAAGGTCTCGGTTTGAT 2177
Db 2230 ACTGTTCTGTTATTCACCTCGATACATAGAGGAGGTTGAAGGTGCTGGTTTGNAT 2289
QY 2178 GATGAGAAATAGTTCTCATGTCTCAATGACTTAGAGAAGAGATTTGGCCAGTCAGCA 2237
Db 2290 GATGAAAAGGCGCTCTTAATGTGCAATATGAGCTGGAGAAGGATTTGACAGTCTGCT 2349
QY 2238 GCATTTGTTGCTCCACTCTGATGAGNAATGTGGTGTCTCCTCAGTCTCCACTCCAGAA 2297
Db 2350 GTTTTTGTGTTCTACCCCTAATGAGAAATGGTGGTTCCTCTTCAGCAACTCCAGAA 2409
QY 2298 TCTCTTTTCAAGAAAGCTATCCATGCTATAAGTTGTGGCTATGAGGACAAAGTCTGAATGG 2357
Db 2410 AACTTTCTCAAGAGGCTATCCATGTCATTAGTTGTGGTTATGAGGATAAGTCAGATTGG 2469
QY 2358 GGAACGTGAGATTGGTTGGATCTATGAGTCTGACAGAAAGATTTCTTACTGGAATCAAG 2417
Db 2470 GGAATGGAGATTGGATGGATCTATGGTCTGTGACAGAAAGATATTCGACTGGGTTCAAA 2529
QY 2418 ATGCAGCAAGAGGCTGGGTTTCAGTCTATTGCATGCCCAAGGCCAGCTTTCAGGCA 2477
Db 2530 ATGCATGGCCGTGGATGGGATCCATTTACTGATGCCCTTAAGCTTCCAGGTTCAAGGGT 2589
QY 2478 TCTGCCCCCATCAATCTTTTCAGATCTCTGAAACAGTCTGCGTGGGCTCTCGGTTCT 2537
Db 2590 TCTGCTCCTATCAATCTTTTCAGATCTCTGAAACAGTCTGCGTGGGCTTTCAGTTCA 2649
QY 2538 GTTGAAATCTTTTCAGCCGGATTCGCCCTTATGATGATGGCTACCGAGGGCGCCCTCAG 2597
Db 2650 GTTGAGATTCTCTCAGTCGGCATTCCTATATGATGATGGTTTACAAATGGAGGCTTAAAA 2709
QY 2598 TTCTCGAGAGATTCCCTTACATACACACACCATTTACCCATACCTCTCTCCGCTT 2657
Db 2710 TTCTGTGAGAGGTTTGGTATGGAACACACCATCTACCCATACCTCTCTCTCTT 2769
QY 2658 CTAGTCTATTGTATTTGCTGCTATCTGTCTCTCACTGGAAAGTTTCATGACCCAGAG 2717
Db 2770 CTCATGATTTGATATTTGCTAGCCGTTGTCTCTTCCACCAACAGTTTATTTCTCTCAG 2829
QY 2718 ATTAGCAACTTGCCAGTATCTGGTTCAATGCGCTCTTCTTTCAATTTTGGCACTGGT 2777
Db 2830 ATTAGTAACATTGCAAGTATATGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
QY 2778 ATCCTTGAGATGAGTGGAGTGTGTGGATTTGAGAGTGGTGGAGGATGAACAGTTTC 2837
Db 2890 ATACTAGAAATGAGTGGAGTGGCTAGGATAGAGCAATGGTGGAGAACAGGACAGTTT 2949
QY 2838 TGGGTCTATTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTTCAGGGTCTTCTTGAAGGTTG 2897
Db 2950 TGGGTCTATTGGTGGATATCCGCTCATTTATTCGCTGTGTTTCAAGGTATCCTCAAAGTC 3009
QY 2898 CTTGCGGATATGACACCAACTTTCAGTCTCACTCAAGGCTAATGACGAAGAGGCGAC 2957
Db 3010 CTTGCGGATATTGACACCAAACTTTCAGATTTACCTTCAAAAGCTTTCAGATGAAGACGGAGAC 3069
QY 2958 TTTGCTGAGCTCTACATGTTTCAAGTGGACGACGCTTCTATCCTCCGACGACATTTTG 3017
Db 3070 TTTGCTGAGCTCTACTTGTTCAAATGGAACAACTTCTGATTCGCAACGAGCTGCTC 3129
QY 3018 ATCAATTAACATGGTTGGTGTGCTGTGGCACTCTCTACGCCATCAACAGTGGTTACCAA 3077
Db 3130 ATTGTAACCTATTGGAGGATTTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCAA 3189
QY 3078 TCATGGGGCGCTCTTTGGGAAGCTCTCTTTGGCTTTCTGGGTGATTTGTTCACTTATAC 3137
Db 3190 TCATGGGGACCACTCTTTGGTAAGTTGTTCTTTGGCTTTCTGGGTGATTTGTTCACTTGTAC 3249
QY 3138 CCATTCCTCAAGGGTCTTATGGGCGAGCAAAACCGCACCGGAGGATTTGTCATCTGCTGG 3197

Db 3250 CTTTCTCAAGGGTTTGTATGGTTCGACAGAACCGGACTCTTACCATTGTTGTGTCTGG 3309
QY 3198 GCTGTCTCTCGCTTCCTATCTTCTTCTTCTGCTGTGGTTCGTTGATCCATCACTACC 3257
Db 3310 TCTGTCTCTTGGCTTCTATCTTCTTCTTCTGTTGTGTGGTTAGGATTTGATCCCTTCTAGC 3369
QY 3258 CTTCTCGTGGCCCAAAATATCCAACTGTGGCATCAACTCACTGCTAGGAA 3305
Db 3370 CGAGTCACTGGCCCGGACATCTTGGAAATGGAATCACTGTTGAGAA 3417

RESULT 11
AAZ99500
ID AAZ99500 standard; DNA; 3725 BP.
XX
AC AAZ99500;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 179..3398
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
(PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84110.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 105-110; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of

CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;

Query Match 38.8%; Score 1408.4; DB 3; Length 3725;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;

```
QY 53 CCGCGCATGACGACGCGGAGCGGAGCCCTCAAGTCCGGGAGGACACGGGGCCGGGAGCGT 112
DB |||||
QY 238 CCGCCACACGCGGACGCGCTGTCCCGGCTAAGCCACGAGAGTGCGAATGGGCGAGGT 297
DB |||||
QY 113 GTGCCAGATCTGCGCCGACGCGCTGGGCACCAACCTTGGACGCGACGCTTTCACCGCTG 172
DB |||||
QY 298 CTGCCAGATTGTGGCGACACTGTGTGGGTTTCAGCCACTGGTGTCTTGTGTGCTG 357
DB |||||
QY 173 CGAGCTGTGCGCTTCCCGTGTGCGGCCCTGTCTACGAGCACGAGCGCAAGAGGGGAC 232
DB |||||
QY 358 CAATGAGTGTGCTTCCCTGTCTGCGGCCCTTGTCTATGAGTACGAGCGCAAGGAGGAA 417
DB |||||
QY 233 CCAGGCTGCTCCAGTGCAGACCAAGTACAAGGCCACAGAGGGAGGCCAGCGATCCG 292
DB |||||
QY 418 CCAATGTGCTCCCTCAGTGCAGACTAGATACAAGAGCAGAAAGTAGCCCTCGAGTTCA 477
DB |||||
QY 293 CGGGGAGGAGGCGACGACACTGATGCCGATGATGCTGACTTCAACTACCCCTGCATC 352
DB |||||
QY 478 TGGTGATGTAGGAGGAGAGTGTGATGACCTGGACNAATCAATATAGCAAG 537
DB |||||
QY 353 TGGCACTGAGGACGAGAGCAGAGATGCTGACAGGATGCGAGTGGCGCATGAACAC 412
DB |||||
QY 538 CAATGGGAAGGGCCGAGAGTGGCAGCTT-----CAAGGAGAT 574
DB |||||
QY 413 CGGGGCGAGTGCATGTGSCCACCACCAAGTATGACAGTGGCGAGATCGSCCTCCAA 472
DB |||||
QY 575 GACCTGATCTGCTTCACTGTGCTGCGCATGACCCACCATCGGATTCACAGCCTTACA 634
DB |||||
QY 473 GTATGACAGTGGAGAGATCCCTAGGGGATAGTCCCTTCACTCCCAACAGCCACGATGC 532
DB |||||
QY 635 AGTGGACA-----ACAGATATC 651
DB |||||
QY 533 AGGAAATCCCTGAGCTTGCCTGATCATCATGATGCTCCCTACGGGGAAACATCAG 592
DB |||||
QY 652 TGGAGAGATCCCTGATCATCCCTGACCGTCACTTATCCGAGTCCAAACATCGAGCTA 711
DB |||||
QY 593 CAGAGCTGCTCCGTTTCCCTATGTAATCATTCACCAATCCGTCAGGGAGTT---CTC 649
DB |||||
QY 712 TGTGATCCAGCGTTCCAGTTCTGTGAGGATTTGTGACCCCTCGAAGGACTTGAATTC 771
DB |||||
QY 650 CGGCAGTATTGGGAATGTTGCTGGAAAGAGAGATTGATGGCTTGGAAATGAAGCAGGA 709
DB |||||
QY 772 CTATGGGCTTAATAGTGTGACTGGAAGGAAAGATTGAGAGCTGGAGGGTTAAACAGGA 831
DB |||||
QY 710 CAAGGTTGCGATTCCTATGACTAATGGGACAGCAATGCTCCCTCTGAGGTCGGGCGAG 769
DB |||||
QY 832 CAAAATATGTGCAAGTGACTAA-----TAAATATCCAGAGGC 870
DB |||||
QY 770 TACTGACATCGATGCATCTACTGATATACAAATGGAAGACGCTTTTACTGAATGATGAAC 829
DB |||||
QY 871 TAGAGGAGACATGGAGGGAGCTGCTCAATGGAAGAGATGCAAAATGTTGATGTC 930
DB |||||
QY 830 TCGCCAGCCTCTATCTAGAAAGTCCCATTTGCTTCCCTCCAAATAAATCCCTACAGAA 889
DB |||||
QY 931 ACGCCTACCTTTGAGCGCGATTTGTCGCAATTTCTCTCAACACAGCTCAACCTTTACCGGAT 990
DB |||||
QY 890 GGTCAATTTGTTGCGGTGTGTTGTTCTAAGCACTTCTGCACTACCGTCTCAAAATCC 949
DB |||||
QY 991 AGTAATCATTTCTCGCTTTATCATCTCTGCTGCTTCTTCTCAATATCGTATCATCTC 1050
DB |||||
QY 950 TGTGCGTAATGCATACCACTGTGGCTTTTATCTGTTATATGTGAGATTGTTGTTGCTTT 1009
DB |||||
QY 1051 AGTCCGTAATGCTTATGAAATGTGGCTAGTATCTGTTATCTGTGAGGCTGTGTTGCTTT 1110
DB |||||
```

```
QY 1010 ATCCTGGATATCTGGATCAGTTCCCGAAGTGTCTTCCAAATCAACCGGAGACCTACTCTTGA 1069
DB |||||
QY 1111 GTCTTGGCTTCTAGATCAGTTTCCCAAAATGATGATCCAAATCAACCGTGAACATATCTCGA 1170
DB |||||
QY 1070 TAGACTGGCTTTAAAGGTATGACCGAGAGGTGAACCGTCTCAGTTGGCTGTGTTGACAT 1129
DB |||||
QY 1171 CAGGCTTGCATTGAGGTATGATGAGAGGGAGAGCCATCAGAGTGGCTCCCATTTGATGT 1230
DB |||||
QY 1130 ATTTGTCAGTACAGTCCGACCCCTTGAAGGAGCCACTATCGTCTACTGCCCAACACTGTGCT 1189
DB |||||
QY 1231 CTTTGTGAGTACAGTGGATCCATTGAAAGGAACTCTCACTGATCAGCCAACTCTGTTTT 1290
DB |||||
QY 1190 ATCCATTCTTGTCTGTGATTATCCGTGGCAAGGTCTCTGCTATGTATGTCTGATGACGG 1249
DB |||||
QY 1291 GTCCATTCTTGTGCTGTGATTTACCTCTGTTGACAAAGTGTCACTGTATGTTCTGTATGAG 1350
DB |||||
QY 1250 AGCTTCAATGTGCTCTTTTGGACGATTTGGCTGTGAGACTTCAGAGTGTGCTAGGAAATGGGT 1309
DB |||||
QY 1351 CTCAGCTATGTGCTGCTTTTGGAGTCTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGT 1410
DB |||||
QY 1310 ACCATTTGTGAAGAGATGACATTTGAACCCAGAGCTCCGAGTGTCTTACTTTTCCGCAAA 1369
DB |||||
QY 1411 TCCCTTTTGAAGAGCAATATTGAACCAAGAGCTCCGAATTTTACTTTTGTCTCAAAA 1470
DB |||||
QY 1370 AATTGATTACCTGAAAGACAAAGTCCAGCCTTCAATTTGTTAAAGACCGCGGGCCATCAA 1429
DB |||||
QY 1471 AATAGATTACTGAGGACAAAATTTCAACTTTCATTTGTTAAAGAAAGACGAGCAATGAA 1530
DB |||||
QY 1430 GAGAGATATGAAGAAATTTAAATCAGGATAAATGCGCTAGTTTCTTAAGGCAATTTGAAAGT 1489
DB |||||
QY 1531 GAGAGATATGAAGAAATTTCAAAATTAAGAAATCAATGCGCTTGTGTTCCAAAGCACGAAAGT 1590
DB |||||
QY 1490 CCCCGGAGAGGATGAGATCATGCAAGATGCACACCATGCGCAGGAACCAATACACAGGA 1549
DB |||||
QY 1591 GCCTGAAAGAGGGTGGACCATGGCTGATGGAATGCTTGGCCTGGGAATTAACCTTAGGGA 1650
DB |||||
QY 1550 TCATCTCTGGAATGATTCAGGTTTTTCCCTGTGTCACAGTGGTGGCTTGATATCTAGAGGTAA 1609
DB |||||
QY 1651 CCATCTCTGGCATGATTCAGGTGTTCTTGGGGCACAGTGGTGGGCTTGACACTGATGAA 1710
DB |||||
QY 1610 TGAGCTCCCGCTTTAGTTTATGTTGCTCTGTGAAAAGCGTCTGGGTTTCAGACCAACAA 1669
DB |||||
QY 1711 TGAATTAACACGCTCTTGTCTATGCTCTCGTGAAGAGAGACCGAGCTTTCAGCATCAAA 1770
DB |||||
QY 1670 GAAGCTGTGTGCGCATGAATGCTTGTGCTGTCTGCTGCTGCTTACTTACTTATGCAATA 1729
DB |||||
QY 1771 GAAGCTGTGTGCGCATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1830
DB |||||
QY 1730 CATGTTGAAATCTTTGATGCTACTACATCAACCAACAGCAAGGCTCTCCGAGAGGCTAT 1789
DB |||||
QY 1831 TCTTCTCAATGTGGATTGTGACCATTAATCAATAGCAGCAAAAGCTCTTAGAGAGCAAT 1890
DB |||||
QY 1790 GTGCTTCTTAATGGATCCAAACCTAGTCCGAAAGTCTGTTATGTGAGTTCACCAAG 1849
DB |||||
QY 1891 GTGCTTCTAATGGATCCAGCTCTAGGAAGGAAATTTGTTATGTACAATTTCCACAAAG 1950
DB |||||
QY 1850 GTTTGATGAGTATGATAGGAATGATGATGATGCAACAGGAACACTCTCTTTTGTGATAT 1909
DB |||||
QY 1951 ATTTGATGGCAATGACTTGTGACGATGATGATGCTAATAGGAACATAGTCTTCTTTGATAT 2010
DB |||||
QY 1910 TAACTTTGAGGGGCTTTGACGGCAATTCAGGACCAAGTTTATGTGGAACTGTTGTGTTTT 1969
DB |||||
QY 2011 CAACATGAAAGGCTTAGATGGCAATTCAGGCTCAGTCTATGTGGGAACAGGATGCTGTTTT 2070
DB |||||
QY 1970 CAACAGAACAGCTATCTATGTTTATGAGCC---CCCAATTAAGCGGAAGAGCCAGGTTTT 2026
DB |||||
QY 2071 CAATAGCAGGCTTTGTTATGGATATGATCTGTTTGTGACTGAAAGCTGATCTGGAACCTAA 2130
DB |||||
QY 2027 CTTGGCATCACTATGTGGGGCAAGGAAGGCAAGCAAGTCAAAAGAAAAGGAGCTCAGA 2086
DB |||||
QY 2131 CATTTGTTGTTAAGAGCTGTGTGTTAGAGGAAGAAAGAAACAGAGTTTATGATGATAG 2190
DB |||||
QY 2087 TAAGAAAAAGTCCGAACAGCATGTGGACAGTTCTGTTTCCAGTATTCAATCTCGAAGACAT 2146
DB |||||
```


2251	CGAGAGGGGTA	TGAAGGT-----TATGAGGA	GAAGGTCA	GTGTATGTCCAGAG	2304
2207	GACCTTAGAGAA	GAGATTTGGCCAGT	CAGCAGCA	TTTGTTCCTCCACTCTGATGAA	2266
2305	GAAATTTGGAAA	ACGCTTTGGTCAGT	CTCAAATCTTCA	TTTGATCCACCTTTATGACTCA	2364
2267	TGTTGGTGTTC	TGACTCTCCACT	CTCAGAAATCTCTTTT	GAAAGAAGCTATCCATGTCTAT	2326
2365	AGTGGGCATAC	CACTTCAACAAAC	CCAGCTTCTACT	GAAAGAAAGCTATCCATGTAT	2424
2327	AAGTCTGGCTAT	GAGGACAAAGTCT	GAAATGGGAA	CTGAGATTTGGTATCTATCGATC	2386
2425	CAGCTGTGGGT	ACGAGACAAAAT	CTGAAATGGGAA	AAGAGATTTGGCTGGATCTATGGTTC	2484
2387	TGTCACAGAAG	ATATTTCTTACT	CGAATCAAGAT	TCGACGAAGAGGTGGCGTTCA	2446
2485	AGTTACAGAGG	ATATTTCTGACT	TGGGTTTAAAT	TGCATGCAAGAGGCTGGCAATCAATCTA	2544
2447	TTGCATGCCAA	AGGCCACGCTTT	TAAGGGATCTGCC	CCCATCAATCTTTTCAAGATCGTCT	2506
2545	CTGCATGCCAC	CAACGACCTTGT	TTCAAGGGTCTTG	CCCAATCAATCTTTCTGATCGTCT	2604
2507	GAACCAAGTGT	CGGGTGGGCTCT	CGGTCTGTTG	AAATCTTTTTCAGCGGATTTGCC	2566
2605	TAACTAGGTGT	CCGTTGGGCTCT	TGGGTCA	GTGGAATTTCTGCTAGCAGACATTTGTC	2664
2567	CTTATCGTAT	TGGCTACGGAG	GGCGCTTCAAGT	TTCTTGGAGAGATTCGCTTACATCAACAC	2626
2665	TATATGATAT	TGGCTCAAT	TGGGCGATTTGA	AGCTTTTGGAGAGGCTGGCTTACATTAACAC	2724
2627	CACCATTTAC	CCCACTAACC	TCTCCCGCTTCT	PAGTCTATTTGTATATTTGCTCTATCTG	2686
2725	CATTGTTTAT	TCCAAATCAC	ATCTGTTCCGCT	TATGTGTCTTCTGCTATCTG	2784
2687	TCTGCTAC	TGGAAAGTTT	ATCATATGCAG	AGATTTAGCAACTTGGCCAGTATCTGTTTCAT	2746
2785	TCTCTTAC	CAATATAATTT	ATCAATCCTG	AGATTTAGTAAATTTTCTTTCAT	2844
2747	TGGCTCTCT	CTTTCAATTTT	CGCCACTGTAT	CTTGCATGATGAGGTGGAGTGGTGTGG	2806
2845	TCTTCTTT	TGCGCTCA	ATTTTCCGA	CTGGTATATTTGGAGCTCAGATGGAGTGGTGTGG	2904
2807	CATTGACG	AGTGGTGGA	GAATGAACAG	TTCTGGGTCAATTTGGAGGTATCTCTGCACATCT	2866
2905	CAATTGAAG	ATTGGTGAG	AAATGAGCAG	TTTGGGTTATTTGGTGGCACCTCTGCCCATCT	2964
2867	GTTTGGCG	CTTTTCA	GGGTCTTCTG	AAAGTGTCTTGGCGGTATGCACACCACTTCATCT	2926
2965	CTTTCGCG	GTGTTCC	AGGGTCTGCTG	AAAGTGTGGGTGGGATTTGATACCAACTTCACAGT	3024
2927	CACCTCAA	AGGCTAATG	ACGAAGA	AGCGACTTTGTCTGAGCTCTACATGTTTCAAGTGGAC	2986
3025	TACCTCAA	GGCATCTG	ATGAGAT	GGCGACTTTGTCTGACTATATGTGTTCAAGTGGAC	3084
2987	GAGCCTTCT	CATCCCTCC	GACGACCA	TTTGTGATCAATTAACATGGTGGTCTCGTGTCTGG	3046
3085	CAGTTTGT	CTCATCCCT	CCGACCA	CTGTCTTGTTCATTAACTGGTGGGATTTGGTGGCAGG	3144
3047	CACCTCTAC	CGCATCA	ACAGTGGTT	TACCAATCATATGGGGGCGGTCTTTTGGGAAGCTCTT	3106
3145	AATTTCTG	TATG	CCCAATTAAC	AGCGCTACCAATCCTGGGGTCCGCTCTTTTGGAAAGCTGTT	3204
3107	CTTTG	CGCTTCTGG	TGATGTCT	CACTTATACCAATTCCTCAAGGGTCTTATGGCAGCA	3166
3205	CTTCTCG	ATCTGGGTG	ATCTCTCAT	CTCTTACCCCTTCTCAAGGGTCTCATGGGACGCA	3264
3167	AAACCG	CACACCG	ACAGTATG	TCATCGTCTGGGTGTCTCTCTCGTCTTCTATCTTCTCCTT	3226
3265	GAACCG	CACG	CAACATCTG	TATCTGTTGGTCCATCTCTTGGGTCTATCTTCTCCTT	3324
3227	GCTGTGG	TTGTTG	ATCCAT	TCATACCC	3258
3325	GCTGTGG	TTGAA	AGATCG	ATCTCTTTCATCTGCC	3356

RESULT 13
 AAZ99506
 ID AAZ99506 standard; DNA; 3725 BP.
 XX AC AAZ99506;
 XX AC AAZ99506;
 DT 03-JUL-2000 (first entry)
 XX AC AAZ99506;
 XX AC AAZ99506;
 DE DE AAZ99506 standard; DNA; 3725 BP.
 XX AC AAZ99506;
 KW Maize; cellulose synthase; stalk quality; stand; si
 transgenic plant; plant breeding marker; ss.
 XX AC AAZ99506;
 OS Zea mays.
 XX AC AAZ99506;
 FH Key Location/Qualifiers
 FT CDS 179..3397
 FT FT /*tag= a
 FT FT /product= "cellulose synthase"
 FT FT /note= "no termination codon given"
 XX AC AAZ99506;
 PN W0200009706-A2.
 XX AC AAZ99506;
 PD 24-FEB-2000.
 XX AC AAZ99506;
 PF 16-AUG-1999; 99WO-US018760.
 XX AC AAZ99506;
 PR 17-AUG-1998; 98US-0096822P.
 XX AC AAZ99506;
 PA (PION-) PIONEER HI-BRED INT INC.
 XX AC AAZ99506;
 PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX AC AAZ99506;
 DR WPI; 2000-224343/19.
 DR P-PSDB; AAY84112.
 XX AC AAZ99506;
 PT New genes which encode maize cellulose synthase pol
 useful for modulating the expression of cellulose s
 to produce transgenic plants expressing the novel p
 XX AC AAZ99506;
 PS Claim 1; Page 121-126; 119pp; English.
 XX AC AAZ99506;
 CC The present sequence encodes a maize cellulose synt
 cellulose synthase can be used for the improvement
 CC improved stand or silage. It also provides an incre
 CC cellulose in the pericarp, hardening the kernel and
 CC handling ability. The sequences are used to produce
 CC seeds expressing the cellulose synthase. The polyn
 CC modulating, preferably increasing, the level of the
 CC cell. The plants are preferably monocots. The polyn
 CC as a probe or primer in the detection quantitation
 CC transcripts. The probes are useful in detecting def
 CC of mRNA in screenings for desired transgenic plant,
 CC mutations in the gene, for monitoring upregulation
 CC changes in enzyme activity in screening assays of c
 CC detection of any number of allelic variants of the
 CC molecular markers in plant breeding programs. The i
 CC of the present invention can also be used for recom
 CC their encoded polypeptides or for use as immunogens
 CC and/or screening of antibodies. The proteins can be
 CC for enzyme agonists or antagonists of enzyme functi
 CC immunogens or antigens to obtain antibodies specifi
 CC with a protein
 XX AC AAZ99506;
 SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U;

Query Match	38.8%	Score 1408.4;	DB 3;	Length 3725;
Best Local Similarity	56.8%	Pred. No. 0;		
Matches 2147;	Conservative	0;	Mismatches 966;	Indels 99;
Gaps	6;			
Qv	53	CCGCGCATGACGCGGACGCGGACGCTTGAATCGGGAGGACACGGGCGCGGGACGT	112	

Db 238 CCGCCACGACGGCGCGCTGTCGGCTAAGCCACGAGAGTGGCAAGT 297
Qy 113 GTGCCAGATCTGCGCGACGCGCTGGGACCACTGGTGGACCGGACGCTTTCAACGCGTG 172
Db 298 CTGCCAGATTTGTGGCGACACTGTGGCGTTTCAGCCACTGGTGATGCTTTGTGTGCGT 357
Qy 173 CGAGCTCTGCGGCTTCCGGTCTCGCGCTCTAGCAGCAGAGCGCAGGAGGCAC 232
Db 358 CAATGAGTGTGCTTCCCTGCTGCGCGCTTGTATGAGTACGAGCGCAAGGAAGGAA 417
Qy 233 CCAGGCTGCTCCAGTGCAGACCAAGTACAAGCGCCACAGAGGAGCCCGAGCTCCG 292
Db 418 CCAATGCTGCCCTCAGTGCAGACTAGATACAAGAGACAGAAAGTAGCCCTCGAGTTCA 477
Qy 293 CGGGAGGAGGCGACGACACTGATGCGGATGATGATGATGATGATGATGATGATGATG 352
Db 478 TGGTGATGATGAGGAGGAAGATGTTGATGACCTGGCAATGAATTCAACTATAAGCAAG 537
Qy 353 TGGCACTGAGGACAGAGCAAGAGATGCTGACAGGATGCGAGCTGGGCGATGAACAC 412
Db 538 CAATGGGAAGGCCAGAGTGGCAGCTT-----CAAGGAGAT 574
Qy 413 CGGGGCGAGTGGCAATGTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAA 472
Db 575 GAGCTGATCTGCTTCACTGCTCGCATGACCCACACCATCGAATTCAGCGCTTACA 634
Qy 473 GTATGACAGTGGAGAGATCCCTAGGAGATAGCTCCCTTCAGTACCACAGCAGATGTC 532
Db 635 AGTGGACA-----ACAGATATC 651
Qy 533 AGGAGAAATCCCTGAGCTTCGCTGATCATCATGATGTCCTACGGGGAACATCAG 592
Db 652 TGGAGAGATCCCTGATGATCCCTGACCGCTCATCTATCGGAGTCCAACTCGAGCTA 711
Qy 593 CAGACGTCTCGCTTCCCTATGTAATCATTCACAAATCCGTCAGGAGATT---CTC 649
Db 712 TGTGTATCCAAAGCGTTCAGGCTTCCTGTGAGGATTTGACCCCTCGAAGGACTTGAATTC 771
Qy 650 CGGCGATTTGGGAATGCTGCTGGAAAGAGAGAGTGTGATGCTGGAAATGAAGCAGGA 709
Db 772 CTATGGGCTTAATAGTGTGACTGGAAGGAAGAGTTGAGAGCTGGAGGGTTAAACAGGA 831
Qy 710 CAAGGGTGCATTCCTCATGACTAATGGAACAAGCATTTGCTCCCTCTGAAGTGGGCGAC 769
Db 832 CAAAAATGTTGCAAGTGACTAA-----TAAATATCCAGAGGC 870
Qy 770 TACTGACATCGATGCTACTACTGAATACAACATGGAAGACGCTTTACTGATGATGAAC 829
Db 871 TAGAGGAGACATGAGGGGAGCTGGCTCAAAATGGAGAAGATATGCAAAATGGTTGATGTC 930
Qy 830 TCGCAGCCTCTATCTAGAAAGTCCCAATGCTTCTCCAAATAAATCCCTACAGAT 889
Db 931 ACGCTACCTTTGAGCGCGCATGTCGCAATTTCTCMAAACGACGCTCAACCTTTACCGGAT 990
Qy 890 GGTCAATGTTCTCGGTTGTGTGTTCTAAGCATCTTCTGCACTACCGTCTCAAAATCC 949
Db 991 AGTAATCATTTCTCGCTTATCATCTGCTGCTTCTTCTTCCAATATGATCAGTATCC 1050
Qy 950 TGTGCGTAATGCTATACCACTGCGGCTTTATCTGTTATATGAGAGATTTGGTTGCTTT 1009
Db 1051 AGTGCCTAATGCTTATGGAATGTGGCTAGTATCTGTTATCTGTTAGGCTCTGTTTGCCTT 1110
Qy 1010 ATCTGATGACTGATCAGTATCCGAGTGGTTCCAAATCAACGGGAGACCTTACCTTGA 1069
Db 1111 GTCTGCTGCTTATAGATCAGTTCCAAATGGTATCCAAATCAACCGTGAGACATATCTCGA 1170
Qy 1070 TAGACTGGCTTTAAGGTATGACCGAGAGGAGTGAACCGTCTCAGTTGGCTGCTGTTGACAT 1129
Db 1171 CAGGCTTGCATTTAGGATATGATAGAGAGGAGAGCCATCAAGCTGCTCCCATTTGATGT 1230
Qy 1130 ATTTGTGATGATGACGCTTTGAAGGAGGAGCCATCTATGCTGCTGCTGCTGCTGCTGCT 1189

Db 1231 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
Qy 1190 ATCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249
Db 1291 GTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350
Qy 1250 AGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
Db 1351 CTCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1410
Qy 1310 ACCATTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTCTTTTACCTTTCCGAGAA 1369
Db 1411 TCCCTTTGTGAAGAGCACAATTTGAACCAAGAGCTCCAGAAATTTTACCTTTGCTCAAAA 1470
Qy 1370 AATTGATTTACTGAAAGACAAAGTCCAGCTTCAATTTGTTTAAAGACCGCGGCGCATGAA 1429
Db 1471 AATAGATTTACTGAAAGACAAATTTCAACCTTCAATTTGTTTAAAGAAAGACGAGCAATGAA 1530
Qy 1430 GAGAGAAATATGAAGAAATTTAAATTCAGGATAAATGCCCTAGTTTCTTAAGGCATTTGAAAGT 1489
Db 1531 GAGAGATATGAAGAAATTTCAAAATAGAAATCAATGCCCTTGTTCGCAAGACAGAAAGT 1590
Qy 1490 CCGCAGGAAGGATGGAATCATGCAAGATGGCACACCATGGCCAGGAACAATACAGGGA 1549
Db 1591 GCCTGAAGAGGGGTGGACCATGGCTGATGAACTGCTTGGCTGGGAAATAACCCCTAGGGA 1650
Qy 1550 TCATCTGGAATGATTTGAGTTCAGGTTTCTTGTGTCACAGTGGTGGCTTGTATCTAGAGGTAA 1609
Db 1651 CCATCTCGGATGATTCAGGTTTCTTGGGCGCACAGTGGTGGCTTGTATCTGATGAGAA 1710
Qy 1610 TGAGCTCCCGCTTTAGTTTATGCTCTCGTGAAGAGCGTCTGGGTTTCAGACACCAAA 1669
Db 1711 TGAATTTACAGCTCTTGTCTATGCTCTCTGTTGAAGAGAGACAGGCTTTTACGATACAA 1770
Qy 1670 GAAGGCTGGTCCCATGAATCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1729
Db 1771 GAAGGCTGGTGAATGAATGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1830
Qy 1730 CATGTTGAATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1789
Db 1831 TCTTCTCAATGTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1890
Qy 1790 GTGCTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1849
Db 1891 GTGCTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1950
Qy 1850 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1909
Db 1951 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2010
Qy 1910 TAACTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 1969
Db 2011 CAACATGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 2070
Qy 1970 CAACAGAAACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2026
Db 2071 CAATAGCAGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2130
Qy 2027 CTTGGCATCACTATGTTGGGCGAAGAGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2086
Db 2131 CATTTGTTTAAAGAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2190
Qy 2087 TAAGAAAAGCTCGAAGCAAGCATGTCAGCTGTCAGTATGATGATGATGATGATGATGATGAT 2146
Db 2191 TCAAGCCGCTTATGAGAGAGACAGATCTTCAAGCTCCCATCTTTTAAATGGAAGACAT 2250
Qy 2147 AGAGGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 2206
Db 2251 CGAGGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 2304
Qy 2207 GAGCTTAGAGAGAGATTTGGCCAGTACAGAGCATTTGTTGGCTTCCCTGCTGATGATGAT 2266
Db 2305 GAAATTTGAGAGAAACGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2364

QY	2267	TGGTGGTGTTCCTCAGTCTCTCCACTCCAGAAATCTCTTTTGAAGAAGCTTATCCATGTCAT	2326
DB	2365	AGGTGGGATACCACTTCAACAAACCCAGCTCTCTACTGAAGGAAGCTATCCATGTTAT	2424
QY	2327	AAGTTGTGGCTATCAGGACAAAGTCTGAATGGGAACCTGAGATTGGTTGGATCTATGGATC	2386
DB	2425	CAGCTGTGGGTACGAGACAAACTGAATGGGGAAGAAGATTGGCTGGATCTATGGTTTC	2484
QY	2387	TGTCACAGAGATATCTCTGATGATCAAGATGCAACGAAGAGGCTGGCGTTCAAGTCTA	2446
DB	2485	AGTTACAGAGATATCTGATCGGGTTTAAAAATGCATGCAAGAGGCTGGCAATCAATCTA	2544
QY	2447	TTGCATGCCAAAGCGCCAGCTTTCAGGGATCTGCCCCATCAATCTTTTCAGATCGTCT	2506
DB	2545	CTGCATGGCACCAAGCACTTGTGTTCAAGGGTCTGCACCAATCAATCTTTCTGATCGTCT	2604
QY	2507	GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGCGCGCATGGCCC	2566
DB	2605	TAATCAGTGTCTCGGTTGGGCTCTTGGCTCAGTGGAAATCTGCTTAGCAGACATTCCTCC	2664
QY	2567	CTTATGCTATGGCTACGGAGGGCCCTCAGTTCCTGGAGAGATTCGTTTACATCAACAC	2626
DB	2665	TATATGCTATGGCTACCAATGGCGGATTTGAAGCTTTTGGAGAGGCTGCTTACATTAACAC	2724
QY	2627	CACCAATTACCCACTAACCTCTCTCCCGCTTCTAGTCTATTGTATATATGCTGCTATCTG	2686
DB	2725	CATTGTTTATCCAAATCACATCTGTTCCGCTTATCGCCCTATTGTGCTTCTCTGCTATCTG	2784
QY	2687	TCTGCTCAGCTGGAAAGTTCATATGCCAGAGATTAGCAACTTGGCCAGATATCTGGTTTCAT	2746
DB	2785	TCCTTCTTACCAATAAATTTATCATCTCTGAGATTAGTAATTTATGCTCGAAATGTTCTTCAT	2844
QY	2747	TGCGCTCTTCTCTTCAATTTTCGCCACTGGTATCTCTTGAGATGAGTGGAGTGGTGTGG	2806
DB	2845	TCCTCTTTTTCGCTCCATTTTCGCCACTGGTATATATGGAGCTCAGATGGAGTGGTGTGG	2904
QY	2807	CATTGACGATGGTGGAGGAATGAACAGTTCCTGGGTCAATGGAGGATATCTTCACACATCT	2866
DB	2905	CATTGAAGATTGGTGGAGAAATCAGCAGTCTTTGGGTTATTGGTGGCACCTCTGCCCATCT	2964
QY	2867	GTTTGGCGTCTTTTCAGGCTCTCTCGAAGTGTCTGCGGTATCGCGGTATCGACACCACTTCAGTGT	2926
DB	2965	CTTCGCGGTGTTTCAGGGTCTGCTGGAAGTGTGGCTGGGATGGATACCACTTCACAGT	3024
QY	2927	CACCTCAAGGCTAATACGAAGAAGCGACTTTTGCTGAGCTCTACATGTTCAAGTGGAC	2986
DB	3025	TACCTCAAGGCACTCTGATGAGGATGGCACTTTGCTGAGCTATGTTTCAAGTGGAC	3084
QY	2987	GAGGCTTCTCATCCCTCCGAGCAATTTGATCATTAAGATGGTGGTGGTCTGGTCTGG	3046
DB	3085	CAGTTTGCTCATCCCTCCGACCACTGTCTTGTCATTTAACTGCTGGGAATGGTGGCAGG	3144
QY	3047	CACCTCTACGCCATCAACAGTGGTTTACCATCATGGGGGCCGCTCTTTTGGGAAGCTCTTT	3106
DB	3145	AAATTCGATAGCCATTAAACGGGCTACCAATCCTGGGGTCCGCTCTTTTGGAAAGCTGTT	3204
QY	3107	CTTTGCGCTTTGGGTGATTGTTTCACTTATACCCATTCCTCAAGGGTCTTATGGGCAGGCA	3166
DB	3205	CTTCTCGATCTGGGTGATTCCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGCAGGCA	3264
QY	3167	AAACCGCACACCGACGATGTCATCGTCTGGGCTGCTCTCTGCTTCTATCTTCTCCTT	3226
DB	3265	GAACCGCACGCCAACATCGTCATCTGGTTCATCTCTCTGGCTCTATCTTCTCCTT	3324
QY	3227	GCTGTGGGTGCTGTTGATCCATCACTACACC	3258
DB	3325	GCTGTGGGTGAAGATCGATCTTTTCATCTCCC	3356

XX	AAA67114;	
AC		
XX		
XX	31-OCT-2000	(first entry)
DT		
XX		
XX	Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.	
DE		
XX		
KW	Eucalyptus grandis; pinus radiata; Monterey pine; modification;	
KW	plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;	
KW	transgenic plant; ds.	
XX		
OS	Pinus radiata.	
XX		
XX	WO200022092-A2.	
XX		
PD	20-APR-2000.	
XX		
XX	08-OCT-1999;	99WO-NZ000169.
PF		
XX		
PR	13-OCT-1998;	98US-00170862.
XX		
PR	11-AUG-1999;	99US-0148426P.
XX		
XX	(GENE-) GENESIS RES & DEV CORP LTD.	
PA	(FLET-) FLETCHER CHALLENGE FORESTS LTD.	
PA		
XX		
XX		
PI	Blocksberg LN;	
XX		
XX	WPI; 2000-339328/29.	
DR		
DR	P-PSDB; AAB16307.	
XX		
XX		
PT	New genes encoding proteins involved in a plant polysaccharide	
PT	biosynthetic pathway, useful for modulating or altering the	
PT	polysaccharide content, composition or structure of the plant.	
XX		
XX		
PS	Claim 1; Page 71-72; 301pp; English.	
XX		
CC	The present invention describes isolated polynucleotides (PN) comprising	
CC	a sequence selected from one of 835 nucleotide sequences given in	
CC	AAA67073 to AAA67907, their (reverse) complements, sequences producing an	
CC	Expectation (E) value of 0.01 or less compared to the 835 sequences,	
CC	sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the	
CC	835 sequences or sequences that are degenerately equivalent or allelic to	
CC	the 835 sequences. The polynucleotides are used to modify the activity of	
CC	a polypeptide involved in a polysaccharide biosynthetic pathway in the	
CC	plant. they are especially used to modulate or alter the polysaccharide	
CC	content, composition or structure of the plant. AAB16268 to AAB16340 are	
CC	proteins encoded by some of the polynucleotide sequence given in the	
CC	present invention	

RESULT 14
AAA67114
ID AAA6

Db 456 GAAGACGTTGATGACATAGAACATGAATTTAATGTGGAGACTCAGCAAAAGAAAACAGGCG 515
Qy 375 AAGATTGCTGACAGGATCGGAGCTGGCGCATGAAACACCGGGGCGAGTGGCAATGTTGGC 434
Db 516 CAGATCACCGAGGGGATGCTCCACGGACGATGAGCTATGGCCGAGGTCCCGACGACGAA 575
Qy 435 CACCCCAAGTATG-ACAGTGCGGAGATCGGCTCTCCA-AGTATGACAGTGGAGAGATCC 492
Db 576 AATTGCGAGATTGCTCATATTCAGAGTTCCTCCGAGATTCTGTACTTGGAAACGGC 635
Qy 493 CTAGGGATACGTCCTTCAGTACCAACAGCCAGATGTCCAGGAGAAATCCCTCGAGCTT 552
Db 636 CACTCGGTTGTAGTGGGAGATTCCAACGTCATACTACGACAGCAACCAATTG----- 689
Qy 553 CGCCTGATCATCATGATGTCCCTACGGGGAACATCAGCAGAGCTGCTCCGTTCCCT 612
Db 690 ---CTTGCCAACCTGCAATGCTGAAGCGTGTGCATCCAAGCTCCGAGCCGGGAGTGA 746
Qy 613 ATGTGAATCATTCACCAAAATCCGTCAAGGGAGTTCTCCGCGAGTATTTGGGAATGTTGCC 672
Db 747 AGGATCATGAGTCCAAACAGGATATTGTTCTTATG--GCTTTGGGAACGTGCTT 804
Qy 673 GGAAGAGAGAGTGTGAGTGTGGGAAATGAAGCAGGACCAAGGTTGCGATTTCCCATGACTA 732
Db 805 GGAAGAGCGAGGCGATGGTTATAAATCGAAGGAAACAAATCAGGCCAGTTGGATATGA 864
Qy 733 ATGGGACAGATTGCTCCCTCTGAAGTCCGGCAGCTACTGACATCGATCATCTACTG 792
Db 865 CGGAAGGAGATATCAATATTAATGGGGGTTTGCACCAATGAGCCCTGAAGATTATATTG 924
Qy 793 AATACAAATCGAAGAGCGCTTACTGAATGATGAAACTCGCCAGCTCTATCTAGAAAAG 852
Db 925 ATCCCGATATG-----CAATGACCGATGAAGCAAGCAGCCACTGTCCCGAAG 975
Qy 853 TCCCAATGCTCTCCAAATAAATCCCTACAGAAATGGTCAATTTGTTCTCGGTTGGTTG 912
Db 976 TGCCAAATTCCTTCAAGCAAAATAAATCCATACCGAATGGTCAATTTGAATTCGATGATAG 1035
Qy 913 TTCTAAGCATCTCCCTGACCTACGCTCTCACAATCTGTGCGTAATGTCATACCCACTGT 972
Db 1036 TGCTGGGTATTTTCTCCGCTATCGCTCTCTGATTCAGTGAAGATGCAATGCGGCTCT 1095
Qy 973 GGCTTTTATCTGTTATATGAGATTTGGTTTGTCTTTATCTCTGGATCTGATCAGTTCC 1032
Db 1096 GGGCCACTTCTATCGTTTGTGAATCTGGTTTGTCTTGTCTCATGATTTCTGATCAGTTTC 1155
Qy 1033 CGAAGTGGTTTCAATCAACCGGAGAGCTTACCTTTGATAGACTGGCTTTAAGGTATGACC 1092
Db 1156 CCAAGTGGTTGCCCTATCAGTCTGTAAGACGTATCTCTGATCGACTGTCAATTAAGGTACGAAC 1215
Qy 1093 GAGAAGGTGAACCTCTCAGTTGCTGCTGTGACATATTTGTGAGTACAGTTCGACCCCT 1152
Db 1216 GAGAAGGGGAACCAATCAATGCTTGACCTGTTGACCTTTTGTGAGTACTGTAGATCCAC 1275
Qy 1153 TGAAGGAGCCACCTATCGTCACTGCCAACACTGTGCTATCOATTTCTGCTGTGATTATC 1212
Db 1276 TGAAGGAGCCCTCTTGGTTACTGCCAATACAGTATTATCAATCCTTTTCAGTACACTACC 1335
Qy 1213 CCGTGGACAAGGCTCTCTGCTATGATCTGATGAGGAGCTTCAATGCTGACTTTTGACG 1272
Db 1336 CTGTAGACAATGTCTCTGTTTATGCTCTGTATGCGGAGCGCTGATGCTTACTTTTGAAT 1395
Qy 1273 CATTTGGCTGAGACTTTCAGATTTCTAGGAATGGGTACCAATTTGTGAAGAGATGATGACA 1332
Db 1396 CTCTCTGTGAGACCTCAGAATTTGCGAGAAATGGGTACCAATTTGCGAAGAAATTCGACA 1455
Qy 1333 TTGAACCCAGAGCTCCCGAGTTTACTTTTTCGCCAGAAATTTGATTACCTGAAAGCAAAAG 1392
Db 1456 TTGAGCTCGGCTCCCGAAATCTATTCTCTCAGAAATTTGACTATCTGAAGGACAAAT 1515
Qy 1393 TCCAGCTTCAATTTGTTAAAGACCGCCGGCCATGAAGAGAGATATGAAGAAATTTAAA 1452
Db 1516 TTCACCCACCTTTTGTCCNAGAGCGCGGTGCCATGAAGAGAGATATGAAGAAATTTCAAG 1575

Qy 1453 TCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAAGTCCCGAGGAAGGATGGATCATGC 1512
Db 1576 TGCSCATCAATTCGGTTGGTTGCAAGGCTCTTAAAGTGCCCAAGGAAGGATGGCAATGC 1635
Qy 1513 AAGATGCAACACCATGGCCAGGAAACAATACACAGGATCATCTCTGGAATGATTCAGGTTT 1572
Db 1636 AAGAGGTACGCTTTGGCTTGGTAAATAATACCCGTGACCATCTCTGGTATGATCCAGTGT 1695
Qy 1573 TCCTTGTGTACAGTGGTGGCTTGATPACTGAGGGTAAATGAGCTCCCGGTTTAGTTTATG 1632
Db 1696 TCTTTGGTTCAGTGGCGGCTCGATACAGAAGCAATGAGCTTCTCTCGGTAGTATATG 1755
Qy 1633 TGTCTCGTGAAGAGCGTCTGGGTTCCAGCACCAACAAGAGGCTGGTCCCATGAATGCC 1692
Db 1756 TTTCTGTGAGAGAGACCTGGTTTCCAGCATACAAAGAGCCGCTGCCCATGAATGCTT 1815
Qy 1693 TTGTTCTGTCTCAGCTGTCTTACTAATGSAACAATCATGTTGAAATCTTGAATGTGATC 1752
Db 1816 TGGTTCCGGTTTCTGCTGTGCTCACCAATGCTCCATTTATGCTGAATGTGATC 1875
Qy 1753 ACTACATCAACAACAGAGCTGTCCGAGAGCTATGTGCTTCTTCTAATGATCCAAACC 1812
Db 1876 ACTACATTAACAATAGCAAGGCAATCAGGGAAGGCATGTGCTTTATGATGGATCCTCAGG 1935
Qy 1813 TAGTCCGCAAGTCTGTTATGTGCAGTTTCCACAAAGGTTTGATGGGATTGATAGGAATG 1872
Db 1936 TTGGGAGAAAGTCTGTTATGTCCOATTTCTCTCAGAGATTCGATGGTATTTGATCGCAATG 1995
Qy 1873 ATCGATATGCAAAACAGGAACACTGTCTTTTGTATATTAACCTTGAGGGGCTTCAACGCCA 1932
Db 1996 ACCGTTACGCCAATCGAAACACCGTATTTCTTTGATATCAACATGAAAGGTCTGGATGGAA 2055
Qy 1933 TTCHAGACACAGTTTATGTGGAACTGGTTGTGTTTTCACAGAACAGCTATCTATGGTT 1992
Db 2056 TTCAAGGGCTGTATATGTGGAACTGGATGTCATGTTTCAAGAACAAGCTCTATATGGGT 2115
Qy 1993 ATGAGCCCCCAATTAAGCGCAAGAACGAGTTTCTTGGGATCACTATGTGGGGCGCAAGA 2052
Db 2116 ATGGGCTCTCCAAAGGCCCAAAAGCTCCCAAGATGGTGACCTGTGATTGT-----C 2166
Qy 2053 AGAAGCAAGCAAGCTCAAAAGAAAGAGCTCAGATAAGAAAGAAAGTCAAGCAAGATGTGG 2112
Db 2167 TCCCTTGTGGGCTCTCGTAAGAAAGTCTCCGAAGAAATAATAGTAGCAAGAAAGTGCAG 2226
Qy 2113 ACAGTTCTGTCTCAGTATTCAATCTCGAAGACATAGAGGAGGTGTTTGAAGGTCTGGGT 2172
Db 2227 GAATCCAGCTCCCGCTACAATCTGACCGGATCGAGGAAGGAGTGAAGGT-----T 2280
Qy 2173 TTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTTAGAAGAGAGATTTGGCCAGT 2232
Db 2281 ATGATGACGAAGAGCATTTGTTGATGAGCCAACTAGACTTCGAGAAGAGTTTGGCCAGT 2340
Qy 2233 CAGCAGCATTTGTTGGCTCCACTCTGATGGAATATGTTGGTGTGTTTCTTCAGTCTCCCACTC 2292
Db 2341 CTTCAGCTTTTGTTCATTCACACTCTGATGGAATGTTGGTGTGTTTCCGCAAAACAGCAAAATC 2400
Qy 2293 CAGAACTCTTTTGAAGAGAGCTATCCATGTCAATAGTTGTTGGCTATGAGGACAGTCTG 2352
Db 2401 CAGCTGAATTTGTAAGAGGCTTATCATGTTCATGCTGTGGGATATGAAGACAAACCGG 2460
Qy 2353 AATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAAAGATATTCTTACTGGAT 2412
Db 2461 AATGGGAAAGAGCTTTGGATGGATCTATGGATCAGTACAGAGGACATTTCTGATGGAT 2520
Qy 2413 TCAAGATGCAAGAGGCTGGCGTTCAAGTCTTATGATGCCCAAGCCGCCAGGTTTCA 2472
Db 2521 TCAAGATGCACTTCGAGGCTGGCGTCCATTTACTGTATGCCCAAAACAGAGCAATTC 2580
Qy 2473 AGGATCTGCCCCCATCAATCTTTTTCAGATGTTGAAACCAAGTCTCGGTGGGCTCTCG 2532
Db 2581 AAGGCTCTGCTCCAATCAATCTATCAGACCGTTTGAACCGAGTGTTCGTTGGGCTTTGG 2640

```
QY 2533 GTTCTGTGAATCTTTTTCAGCGGCAATGCCCTTATGATGCTACGGAGGGGCC 2592
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2641 GATCAGTAGAAATTTTCATGAGCAGACATTCGCCAACTCTGGTATGGCTATGGGGAGGTC 2700
QY 2593 TCAAGTTCCTGGAGAGATTCCGTTTACATACACACCAATTTACCCCACTACTCTCTCC 2652
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2701 TGAATGGCTTGAAGATTTGCCCTATATCAACACCAATTTGCTATCCATTACCTCTCTTC 2760
QY 2653 CGCTTCTAGCTATTGCTATATGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2712
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2761 CACTCATGCTATTGACACACTTCCAGCGGTCAGTTTGCCTCACTGCGCAAAATTTGTATCC 2820
QY 2713 CAGAGATTAGCAACTTGGCCAGTATCTGGTTCAATGCGGCTCTCTTTCAATTTGCGCA 2772
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2821 CTGAGATCAGTACTTTTGCAGTCTATTTTAAAGCTCTTTTCACTCTCAATTTTGGCA 2880
QY 2773 CTGCTATCTTGATGAGTGGAGTGGTGTGGCATGAGTGGTGGTGGAGGAATGAAC 2832
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2881 CTGCTATCTTGAATGAGTGGAGTGGTGGATGAGCAATGAAGAAATGCTGGCAAAATGAAC 2940
QY 2833 AGTTCTGGGTCATTTGAGGATCTCTGACATCTGTTTGGCGTCTTTGAGGCTTTCTGA 2892
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2941 AGTTCTGGGTCATTTGAGGAGGTTTCTGACATTTTTCAGATTATTCAGAGTCTGCTCA 3000
QY 2893 AGGTGCTTGCCTGATCGACCACTTCACTGTCTCACTCAAAAGGCTAATGACGAAGAG 2952
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3001 AGGTACTGGCAGGCAATGTATACAAATTTTCAAGTCACTGCGCAAGGC---ATCAGATGACG 3057
QY 2953 GCACATTGCTGAGCTCTACATGTTTAAAGTGGAGCAGCGTCTCATCCCTCGACGACCA 3012
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3058 GTGAGTTTGGGGAACCTGTATGCAATTCAAATGGACACACCTCTCTCTCTTCTTCTCAACCC 3117
QY 3013 TTTTGTATCATTAACATGTTGGTGTGCTGCTGCGACCTCTTACGCAATCAAGTGGTT 3072
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3118 TGTGTGTCATCACTTGTGGGGTGTGTTGCGGTAGCAGATGCAATCAACATGGAT 3177
QY 3073 ACCAATCATGGGGCGCGCTTTTGGGAAGCTCTTTTGGCTTCTGGGTCAATGTTCACT 3132
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3178 TTCAGTCAATGGGTCTCTCTTGGGTAAAGTTTCTTTGATTTCTGGGTCAATTTGTGACC 3237
QY 3133 TATACCATCTCTCAAGGTTCTTATGGGAGGCAAAACCGCAACCGCAAGTGTGATCG 3192
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3238 TGTATCTTTCTCTCAAGGTTCTCATGGGAGGCAAGAAACACACCCATCGTGTGA 3297
QY 3193 TCTGGGCTGTCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3252
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3298 TTTGTCAATCTGCTGGCATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3357
QY 3253 CTACCGTCTCGCTGGGCCCAATATCAAACTGTGGCATCAACTGCT 3300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3358 TGAAGTAAAGTTAAAGGCCAGATCTAAACAATGTGGCATCAACTGCT 3405
```

RESULT 15

AAZ99530
ID AAZ99530 standard; DNA; 3753 BP.

XX AC AAZ99530;

XX 03-JUL-2000 (first entry)

DT DNA encoding a maize cellulose synthase.

DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX OS

XX Key

XX CDS

Location/Qualifiers

184..3405

/tag= a

/product= "cellulose synthase"

/trans_except= (pos: 916..918, aa: Xaa)

/note= "no termination codon given; Xaa is an unspecified amino acid"

WO200009706-A2.

24-FEB-2000.

16-AUG-1999; 99WO-US018760.

17-AUG-1998; 98US-0096822P.

(PTON-) PIONEER HI-BRED INT INC.

Dhugga KS, Helenjaris TG, Bowen BA, Wang X;

WPI; 2000-224343/19.

P-PSDB; AAY84120.

New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein.

Claim 1; Page 184-188; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein

Sequence 3753 BP; 944 A; 805 C; 965 G; 1038 T; 0 U; 1 Other;

Query Match 37.9%; Score 1373; DB 3; Length 3753;
Best Local Similarity 66.3%; Pred. No. 0;
Matches 2130; Conservative 0; Mismatches 986; Indels 96; Gaps 7;

```
QY 53 CCGCGCCATGGACGGCGAGCGGACGCCCTGAAGTCCGGGAGCACGCGGGCGGGGACGT 112
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
243 CCGCCACAGCGGAGTGTCCGGCTCGGCTAGCCCAAGAGTGCAGATGCAGGT 302
QY 113 GTGCCAGATCTGCGCGACGGCTGGGACACAGTTGGACGGCAGCTTTTCAACCCCTG 172
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
303 CTGCCAGATTTGCGGTGACTCTGTGGTGTTCAGCCACTGGTGTGTTCTTTTTCCTG 362
QY 173 CGACGCTCTCCCGTTCCTCCCGTCTCCCGCTCTCCCGCTCTCCCGCTCTCCCGCT 232
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
363 CAATGAGTGTCTTCCCTGTCTGCGGCCCATGCTATGATGATGACGCAAGAGGGGAA 422
QY 233 CCAGGCTCTCCCTCCAGTGCAGACCAAGTACAAAGCCACAGAGGAGGCCAGCGATCG 292
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
423 CCAATGCTGCCCGCCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTCA 482
QY 293 CGGGGAGGAGCGACGACACTGTATGCCGATGATGATGATGATGATGATGATGATGAT 352
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
483 TGGTGTATGAGGATGAGGAAGATGTTGATGACCTAGACATGAATTAATCACTAC----- 534
QY 353 TGGCACTGAGGACGACGACGAAGATTGCTACAGGATGCGGAGTGGCGGATGACAC 412
```

[illegible]

QY	2567	CTTATGGTATGCTAGGAGCGCGCTCAAGTTCCTGGAGAGATTGGCTTACATCAACAC	2626
Db	2673	TATCTGGTATGGTTTAAATGAGAGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACAC	2732
QY	2627	CACCAATTTACCCACTAACTCTCTCCGCTTCTAGTCTATTGTATATTGCTTGCCTGCTATCTG	2686
Db	2733	TATTGTATATCCAAATCACAATCCATTCGCTTATTGCCTATTGTGTCTCCCGCTATCTG	2792
QY	2687	TCCTGCTACTGGAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCA	2746
Db	2793	CCCTCCCTTACCAATAAATTTATCATCTCTGAGATTAGCAATTATGCTGGGATGTTCTTCAT	2852
QY	2747	TCGGCTCTTCCCTTTCAATTTTCCCACTCGTATCCTTGAGATCAGGTGAGTGGTGTGG	2806
Db	2853	TCCTCTTTTTCGGCTCCCAATTTTCCCACTCGTATATTGGAGCTTAGATGGAGTGGTGTGG	2912
QY	2807	CATTGACGAGTGGTGAGGAATGAACAGTCTCTGGGTCAATTGGAGGTATCTCTGCACATCT	2866
Db	2913	CATTGAGATTGGTGAGGAATGAGCAGTTTGGGTATTGGTGGCACCTCTGCCCACTCT	2972
QY	2867	GTTCGCGTCTTTCAGGGTCTCTGAAGGTGCTTGCCGGTATCGACACCAACTTCACTGT	2926
Db	2973	CTTCGAGTGTTCAGGGTCTGTGAAGTGTGGCTGGGATTGATACCAACTTCACAGT	3032
QY	2927	CACCTCAAGGCTAATGACGAGAGGCGACTTGTGCTGAGCTCTACATGTTCAAGTGGAC	2986
Db	3033	TACCTCAAGGCTAATGATGAGATGGGACTTGTGCTGAGCTATATGTTTCAAGTGGAC	3092
QY	2987	GACGTTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGGTGTGTGTTGCTGG	3046
Db	3093	CAGTTTGTCTCATCTCCGACCACTGTTCTTGTCAATTAACTGGTGGATGGTGGCAGG	3152
QY	3047	CACCTCTACGCCAACAACAGTGGTTTACCAATCANGGGGCGCTCTTTGGGAAGCTCTT	3106
Db	3153	AATTTCTTATGCCATTAAACAGTGGCTACCAATCCTGGGTCCGCTCTTTGGAAAGCTGT	3212
QY	3107	CTTTGCCCTTCTGGGTGATTGTTCACTTATACCAATTCCTCAAGGGTCTTATGGGCAAGCA	3166
Db	3213	CTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAAGCA	3272
QY	3167	AAACGCGACACGAGATTGTCATCGTCTGGGTGTCCTCCTCGCTTCTATCTTCTCCTT	3226
Db	3273	GAACGCGACACCAACAATCGTCAATTGCTGGTCCATCCCTTCTGCACTATCTTCTCCTT	3332
QY	3227	GCTGGGGTTCGTGTTGATCCCAATCACTACC	3258
Db	3333	GCTGGGGTGAAGATCGATCCCTTTCATCTCCC	3364

Search completed: August 23, 2004, 01:39:31
Job time : 1325 secs

2478	Qy	TC	TCTGCCCCCAATCAATCTTTTCAGATCGTCTGAA	CCAAAGTGTCTGGTGGGCTCTCGGTTCT	2533
2590	Db	TC	TGTGCTCCTATCAATCTTTTCAGATCGTCTGAA	CCAAAGTGTCTGGTGGGCTTTAGGTTCA	2649
2538	Qy	GT	TGAAATCTTTTCACCGGCATTCGCCCTTATGGT	TATGGCTACGGAGGGCGCTCAAG	2597
2650	Db	GT	TGAGATCTCTTCAGTCGSCATTTGCTATATGGT	TATGGTTACATGGAGGCTAAAA	2709
2598	Qy	TT	CTGTGAGAGATTTCGGTTTACACAA	CACACACATTTTACCACATTAACCTCTCTCCCGCTT	2657
2710	Db	TTT	CTTGAGAGGTTTGGGTATGTGAAC	CACCACCATCTACGCTCATCACTCCATTCCTCTT	2769
2658	Qy	CT	AGTCTATTGTATATTTCGCTGCTCATCTGTCTG	CTCACTGMAAGTTTCATCAGCCAGAG	2717
2770	Db	CT	CAATGATTGTACATTTGCTAGCCGTTTGCTCTT	TTCACCAACACAGTTATTATTTCCTCAG	2829
2718	Qy	AT	TAGCAACTTGGCCAGTATCTGGTTTCATTTGCG	CTCTTCCTTTTCAATTTTCGCCACTGGT	2777
2830	Db	AT	TAGTAACATTCGAAGTATATGGTTTCTGTCT	CTCTTTCTTCCATTTTCGCCACGGGT	2889
2778	Qy	AT	CCCTGAGATGAGGTGGAGTGGTTTGGCATTTGA	CGAGTGGTGGAGGAATGAACAGTTC	2837
2890	Db	ATA	CTAGAAATAGAGTGGAGTGGCGTAGGCATAGAC	GGAATGGTGGAGAAACGACGACAGTTT	2949
2838	Qy	TGG	GTCACTCGAGGTATCTCTGCACATCTGTTTGC	CGCTCTTTTCAGGGTCTTCTGAAGGTG	2897
2950	Db	TGG	GTCAATGGTGGAGTATCGCTCATTTATTTCGT	GTGTTTCAAGTATCTCTCAAGTCT	3009
2898	Qy	CT	TGCCGGTATCGACACAACTTCACCTGTCACTG	TCACTCAAGGCTAATGACGAAGAGCGCAC	2957
3010	Db	CT	TGCCGGTATTGACACAAACTTCACAGTTTACCT	CAAAAGCTTCAGATGAAGCGGAGAC	3069
2958	Qy	TT	TGCTGAGCTCTACATGTTCAAGTGAACGACGCT	TCTCATTCCTCCGACGACCATTTTG	3017
3070	Db	TT	TGCTGAGCTCTACTTTGTTCAAA	TGACAAACATCTCTGATTCGCGCAACGAAGCTGCTC	3129
3018	Qy	AT	CATTAAACATGGTTGGTGTCTGTTGGCACCTCT	TACGCCATCAACAGTGGTTACCAA	3077
3130	Db	ATT	GTAACCTTAGTGGAGTTGTTGACGAGTCTCTT	ATGCTATCAACAGTGGATACCAA	3189
3078	Qy	TC	TGGGGGCGCTCTTTTGGGAAGCTCTTCTTTG	CGCTTCTGGGTGATTTGTTCACTTATAC	3137
3190	Db	TC	TGGGGACCACTCTTTTGGTAACTTCTTCTT	TGSCCTCTGGGTGATTTGTTCACTTGTAC	3249
3138	Qy	GC	ATTCCTCAAGGCTTATGGGACAGGCAAAACCG	ACACCGACATGTCATCGTCTGG	3197
3250	Db	CC	TTTCTCCACAGGCTTTGATGGGTCGACAGAAC	CGGACTCTCTACCATTTGTTGGTCTGG	3309
3198	Qy	GC	TGCTCTCTCGCTTCTATCTTCTCTCTGCTGTG	GGGTTGCTGTTGTATCAATTCACCTACC	3257
3310	Db	TC	TGTTCTCTTGGCTTCTATCTTCTCGTTGTTG	TGGGTAGGATTGATCCCTTCACTAGC	3369
3258	Qy	CG	CTCGCTGGCCCCAAATATCCAAACCTGTGGCA	TCAACTGCTTAGGAA	3305
3370	Db	CG	AGTCTACGTGCCCGGACATCTCGAATGTGGAAT	CAACTGTTGAGAA	3417

RESULT 2

US-09-221-013A-5
; Sequence 5, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699

869 AATATCGTACACTCACCCCTGTGAATAATGCATATCTTTGTTGGTTGACCTCGGTATCT 928
991 GTGAGATTGTTGTTGTTTATCTCTGGATACCTGGATACCTCCGAAAGTGGTTTCAATCA 1050
929 GTGAGATCGTTGTTGCAATTTCTTGGCTTCTTGATCAGTTTCCCAATGGTACCCCATTA 988
1051 ACCGGGAGACCTACTCTGATGAGACTGGCTTTTAAAGTATGACCGAGAAGGTGAACCGTCTC 1110
989 ACAGGGAGACTTATCTTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAACCATCAC 1048
1111 AGTTGGCTGCTGTTGACATATTTGTGACGTACAGTCGACCCCTTGAAGGAGCCACCTATCG 1170
1049 AGTCGTTCTGTTGATGTTGTTGTAGTACAGTGGACCCATTTGAAGAGGCTCCCTTG 1108
1171 TCATGTCACACACTGTGCTATCCATCTTTGTTGTTGATATCCCGTGGACAAAGTCTCT 1230
1109 TTACAGCAAAACACAGTTCTCTCGATCTTCTGTGGACTACCCGGTAGATAAAGTAGCCT 1168
1231 GCTATGATCTGATGACGAGCTTCAATGCTGACTTTTGAAGCTATGCGCTGCTGAGACTTCAG 1290
1169 GTTATGTTTCAGATGATGGTTGAGCTATGCTTACCTTTTGAATCCCTTTCTGAACCGCTG 1228
1291 AGTTTGTGAGAAATGGGTACCATTTGTGAAGAAGTATGACATTTGAACCCAGAGCTCCCG 1350
1229 AGTTTGCAAGAAATGGGTACCATTTTGCAGAAATTTCAACATTTGAACCTAGGGCCCTG 1288
1351 AGTTTACTTTTGCAGAAATTTGATTAACCTGGAAGACAAAGTCCAGCCCTTCAATTTGTTA 1410
1289 AATTCTAATTTTGCAGAAATAGATTTACTTTGAAGGACAAAGTCCAAACCGTCTTTTGTGA 1348
1411 AAGACCGCGGCGCATGAAGAGAGATATGAAGATTTAAATCAGGATTAATGCCCTAG 1470
1349 AAGAGCGACGAGCTATGAAGAGAGAGATGAAGATTTAAAGTAGGATAAATGCTCTTG 1408
1471 TTTCTAAGGCATTTGAAGTCCCGAGGAGGATGATCATGCAAGATGGCACACCATGGC 1530
1409 TTGCCAAGACACAGAAATCCCTGAAGAGGCTGGACATGCAAGATGCTACCTCCCTGC 1468
1531 CAGGAAACAATACAGGAGATCATCTGGAATGATTTCAGGTTTTCCTGGTCAAGTGTG 1590
1469 CTGGTAAACAACACTAGAGATCATCTGGAATGATACAGGTGTTCTTAGGCCATATAGTGGG 1528
1591 GCCTGATACAGGATTAATGAGCTCCCGTTTATGTTTATGTTGCTGCTGAAAGCGTC 1650
1529 GTCTGGATACCGATGGAATGAGCTGCTAGACTCATCTATGTTTCTCGTGAAGAGCGGC 1588
1651 CTGGTTCCAGCACCAAGAGGCTGGTGGCATGAATGCCCTTGTCTGCTCTCAGCTG 1710
1589 CTGGATTTCAACACCACCAAAAGGCTGGAGCTATGAATGCATTCGATCGTATCTGCTG 1648
1711 TCCTTACTAATGGACAAATACATGTTGAATCTTGATTTGATGATCACTACATCAACACAGCA 1770
1649 TTTCTAACCAATGGAGCATATCTTTTGAACGTTGATTTGATCAATTAATTAACAGTA 1708
1771 AGGCTGTCGAGAGACTATGCTTCTTAAAGTATCCAAACCTAGTCCGCAAGTCTGTT 1830
1709 AGGCTATTAAAGAGCTATGTTTCAATGATGGACCCCGCTATTGGAAGAAGTCTGCT 1768
1831 ATGTGCACTTCCCAAAAGGTTTGAATGGGATTTGATAGGAATGATCGATATGCAACACAGGA 1890
1769 ATGTCCAGTTCCTCAACGTTTGAACGTTTGAATTTGACGATTCGATTCGATGCAACAGGA 1828
1891 ACACGTCTCTTTTATGATTAATCTTGAAGGCTTTGAGGCTTTGAGGCTTCAAGGACCAAGTTATG 1950
1829 ATAATGCTCTTTTTCGATTAATCAATGAAGGGGTGATGGATTCAGGGTCCAGTATATG 1889
1951 TGGAACTCGTGTGTTTTCACAGAAACAGCTATCTATGTTTATGAGCC-----CCCA 2003
1889 TGGTACTGTTGTTTAAATAGGACGCTCTATATGGGTATGATCCTGTTTGAACGG 1948
2004 ATTAAGGCAAGAACCCAGGTTTCTTGGCATCA--CTATGTTGGGGGCAAGAAAGGCA 2061

1949 AAGAAGATTTAGAACCAATATTTATTTGTCAAGAGCTGTTTGGGGTCAAGGAAGAAGGTA 2008
2062 GCAAGTCAAGAAAAGAGCTCAGATAAGAAAAGTCGAAACAAGCATGTGACAGTTCGTG 2121
2009 AAAGTAGCAAGAAGTATAACTACGAAAAGAGAGAGGCATCAACAGAAGTACTCCAATG 2068
2122 TTCCAGTATTCAATCTCGAAGACATAGAGAGGGTGTGGAAGGTGCTGGGTTTGATGATG 2181
2069 CTCACATTTTCAATATGAGGACATCGATGAGGGTGTGGAAGGT-----TATGATGATG 2122
2182 AGAATCAGTTCCTCATGCTCTCAAAATGAGCTTTAGAGAAGAGATTTGGCCAGTCAGCAGAT 2241
2123 AGAGTCTATTTCTAATGTGCCAGAGGAGTGTAGAGAAGCGTTTGGTCAAGTCGCGGTAT 2182
2242 TTGTTGCCCTCACHCTCTGATGGAATATGTTGTTGTTCTCAGTCTCTCACHCTCAGAACTC 2301
2183 TTATTTGGGCAACCTTCAATGAACAAGCGCGCATTTCCACCAACAACCAATCCCGTACTC 2242
2302 TTTTGAAGAAGACTATCCATGTCAATAGTTCTGCTATGAGCAAGAAGTCTGAATGGGAA 2361
2243 TTCTGAAGGAGCTTATTCATGTTATAGCTGTGTTAGAGAACAAGTCTGAATGGGCA 2302
2362 CTGAGATTGTTGGATCTATGATCTGTCAAGAGATATTTCTTACTGATTTCAAGATGC 2421
2303 AAGAGATTGTTGGATCTATGTTCCGTGACGGAAGATATTTCTTACTGGGTTCAAGATGC 2362
2422 ACGCAAGAGGCTGGGTTTCACTTATTCATGTCGCAAGCGCCAGCTTTCAAGGATCTG 2481
2363 ATGCCCGGGTGGATATCGATCTACTGCAATCTCCACGCCCTGCGTTCAAGGATCTG 2422
2482 CCCCCTCAATCTTTTCAGATCTGTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTG 2541
2423 CACCAATCAATCTTTCTGATGCTGTTGAACCAAGTCTTTCGATGGGCTTTGGGACTATCG 2482
2542 AAATCTTTTTCAGCGGCAATTTGCCCTTATGTTATGCTACGAGGGGCGCTCAAGTTCC 2601
2483 AGATCTTCTTAGCAGACATTTCTCTATCTGATGTTTACCATGGAAGTGTGAGACTTT 2542
2602 TGGAGAGATTGCTTTATCATCAACACCACTTATCCCACTAACCTCTCTCCGGTCTTAG 2661
2543 TGGAGAGATCGCTTATATCAACACCACTGCTTATCTCTTATTCATCCACTCTCTTATG 2602
2662 TCTATTGTATATGCTGCTATCTGTCTGCTCACTGGAAGTTCATCATGCGCAGAGATTA 2721
2603 GGTATTGTATTCTTCCCGCTTTTGTCTCATCAACGACAGATTCAATATACCCGAGATA 2662
2722 GCAACTGGCGAGTATCTGTTCAATGCGCTCTTCTCTTCAATTTTCCGCACTGATATCC 2781
2663 GCAACTACGCGAGTATTTGGTTCTTCTTCTCTCTCAATTTGCTGTGACTGGAATCC 2722
2782 TTGAGATGAGTGGAGTGGTGGCATTTGAGAGTGGTGGAGGAAAGAAAGTCTCTCGG 2841
2723 TGGAGCTGAGATGGAGGCGGTGAGCATTTGAGGATTTGGTGGAGAACGAGCAGTTCTCGG 2782
2842 TCATTGGAGGTATCTCTGACATCTGTTTGGCTCTTTTCCAGGCTCTTCTGAAAGTGTCTG 2901
2783 TCATTGGTGGACATCCGCCCATCTTTTGTCTCTTCCAAAGTCTACTTAAAGTGTCTG 2842
2902 CCGGTATGCAACCACTTCTGCTACCTCAAGGCTTAATGACGAGAAGGCGACTTTG 2961
2843 CTGGTATCGAACCACTTCAACCGTTTACATCTAAAGCCACAGAGAAAGATGGGATTTG 2902
2962 CTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCTCCCTCCGACGACCAATTTTATCA 3021
2903 CAGAACTCTACATCTTCAATGGACAGCTCTTCTCATTTCCACCAACCGCTCTACTTG 2962
3022 TTAACATGTTGGTGTCTGCTGGCACCTCTTACGCGCATCAACAGTGGTGTACCAATCAT 3081
2963 TGAACCTCATAGGCATTTGCTGGTGTCTCTTATGCTGTAAACAGTGGTACCAGTCTG 3022
3082 GGGGGCGCTCTTTTGGGAGCTCTTCTTGGCTCTCTGGTGTGTTTCACTTATACCAT 3141
3023 GGGGTCCGCTTTTTCGGGAGCTCTTCTTCCGCTTATGGGTTATTTGCCCATCTCTACCTT 3082


```
Qy 1651 CTGGTCTCCAGCACCAAGAGGCTGGTGCATGAATGCCCTGTGTTCTGCTCAGCTG 1710
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1659 CTGGATTCAACACCAACAAAGAGCTGGAGCTATGAATGCATTTGATCGTGTATCTGTTG 1718
Qy 1711 TCCTTACTAATGGCAATACATATGTTGAATCTTGATTTGATGATCACTATCAATCAACACAGCA 1770
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1719 TTCTTACCAATTTGGAGCATATCTTTTGAACGTTGGATTTGATCATTTACTTTAATAACAGTA 1778
Qy 1771 AGGCTGTCCGAGAAGCTATGCTTCTTAATGGATCCAAACCTAGGTCGCGAAGTCTGTT 1830
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1779 AGGCTATTAAAGAAGCTATGTTTTCATGATGGACCCGGCTATTGGAAGAAGTGTGCTGCT 1838
Qy 1831 ATGTGAGTTCCCAACAAAGTTTGTATGGGATTTGATAGGAATGATGATATCAATCAACACAGGA 1890
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1839 ATGTCCAGTTCCCTCAACGTTTTCAGCGTATTGATTTTGCAGATCGATATCCCAACAGGA 1898
Qy 1891 ACAGTGTCTTTTGTATTAATTAACCTTGAAGGCGCTTGACGGCATTTCAAGGACAGTTTATG 1950
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1899 ATATAGTCTTTTTCGATATTAAACATGAAGGGGTTGGATGATCCAGGGTCCAGTATATG 1958
Qy 1951 TGGAACTGGTGTGTTTTCACAGAACAGCTATCTATGTTATGAGCC-----CCCA 2003
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1959 TGGTACTGTTGTTTGTATAGGAGGCTCTATATGGGTATGATCTGCTTTTTCAGG 2018
Qy 2004 ATTAAGCGAAGAACCGAGGTTTCTTGCCATCA--CTATGTGGGGCGAAGAAAGGCAA 2061
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2019 AAGAAGATTAAAGAACCAATATTATTCAAGAGCTGTGGCGGTCAAGGAAGAAAGTA 2078
Qy 2062 GCAAGTCAAGAAAAGGAGCTCAGATGAAGAAAAGTCAACAAAGCATGTGACAGATTTCTG 2121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2079 AAAGTAGCAAGAATATAACTTACCAAGAGAGGAGGAGCATCAACAGAAAGTCACTCCAATG 2138
Qy 2122 TTCCAGTATTCAATCTCGAAGACATAGAGGAGGTTGTTGAAGGTGCTGGGTTTCATGATG 2181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2139 CTCACATTTCAATATGAGAGACATCGATGAGGGTTTGAAGT-----TATGATGATG 2192
Qy 2182 AGAAATCAGTTCTCATGCTCAAAATGAGCTTTAGAGAAGAGATTTGGCCAGTCAGACAGAT 2241
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2193 AGAGTCTATTCTAATCTCCAGAGGAGTGTAGAGAAGCGTTTGGTCAGTCGCGGTAT 2252
Qy 2242 TTGTTGCTCCACTCTGATGAATATGGTGTGTTCTCAGTCTTCCACTCCAGAAATCTC 2301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2253 TTAATGCGGCAACCTTCATGGAACAAGCGCGCATTTCCACCACCAACCAATCCCGCTACTC 2312
Qy 2302 TTTTGAAGAAGCTATCCATGTCTATAAGTTGTGCTATGAGCAAGTCTGAATGGGAA 2361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2313 TTCTGAAGAGGCTATTCAATGTTATAGTGTGTTACGAGACAAAGTGAATGGGCA 2372
Qy 2362 CTGAGATTGGTATGATCTATGATCTGTCAAGAGATATTCTTACTGGATTCAGATGC 2421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2373 AAGAGATTGGTTGGATCTATGGTTCCGTGACGGAAGATATTCTTACTGGGTTCAAGATGC 2432
Qy 2422 ACGCAAGAGCTGCGTTTCACTGCTATTGTCATGCCCAAGCGCCAGCTTCAAGGGATCTG 2481
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2433 ATGCCCGGGTTGGATATCGATCTACTGCAATCTCCACGCGCTCGGTTCAAGGGATCTG 2492
Qy 2482 CCCCCTCAATCTTTTCAGATCGTCTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTG 2541
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2493 CAACCAATCAATCTTTCTGATCGTTTGAACCAAGTCTTCGATGGGCTTTGGATCTATCG 2552
Qy 2542 AAATCTTTTTCAGCGGATTTGCCCTTATGATGATGATGATGATGATGATGATGATGATG 2601
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2553 AGATCTTCTTAGCAGACATTTGCTATCTGTTATGTTTACATGGAAGGTTGAGACTTT 2612
Qy 2602 TGGAGATTTCGTTTACATCAACACCACTTTTACCCACTAACTCTCTCCCGTTCTAG 2661
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2613 TGGAGAGATTCGCTTTATTAACACCATCTGCTATCTATTTACATCATCTCCCTTTATG 2672
Qy 2662 TCTATTGATTAATTCCTGCTATCTGCTGCTCACTGGAAAAGTTTCATCATGCCAGAGATTA 2721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2673 CGTATTGATTTCTCCCGCTTTTGTCTCATCCCGACAGATTTCATCATACCCGAGATA 2732
Qy 2722 GCAACTTGGCCAGTATCTGGTTTATTCGGCTCTTCTCTTCAATTTTTCGCCACTGGATCC 2781
```

```
Db 2733 GCACTAGCGAGTATTGGTTTCTTCTACTCTCTCATCTCAATTCGCTGACTGAAATCC 2792
Qy 2782 TTGAGATGAGGTGAGTGGTGTGGCATTTGACGAGTGGTGGAGAAAGAAAGTTCTCTGG 2841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2793 TGGAGCTGAGATGGAGCGGTGTGAGCATTTGAGGATTTGGTGGAGAAACGAGCAGTTCTGG 2852
Qy 2842 TCATTGGAGGTATCTCTGCACATCTGTTTTCGCTTTTTCAGGGTCTTCTGAAGGTGCTTG 2901
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2853 TCATTGGGACATCCGCCCATCTTTTGTGCTCTTCCAAGGTCTTCTTAAGGTCTCTTG 2912
Qy 2902 CCGGTATCGACACCAACTTCACTGTCACTCAACCTCAAGGCTTAATGACGAAGAAGCGCACTTTG 2961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2913 CTGTATCGACACCAACTTCACTGTCACTCAACCTTCAAGGCTTAATGACGAAGAAGCGGATTTG 2972
Qy 2962 CTGAGCTCTACATGTTCAAGTGGAGGAGCTTCTCATCTCCCTCCAGACGACATTTTGATCA 3021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2973 CAGAACTCTACATCTTCAAAATGGACAGCTCTTCTCAATTCACCAACCAACCGTCTCTACTTG 3032
Qy 3022 TTAACATGGTGTGTTGCTGGCACCTCTACGCGCATCAACAGTGGTTTACCAATCAT 3081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3033 TGAACCTCATAGGATTTGTGGCTGTCTCTTATGCTGTAAACAGTGGCTACCAAGTCTG 3092
Qy 3082 GGGGGCGCTCTTTTGGAAAGCTCTTCTTTCGCTTCTGCGTGAATTTTCACTTATACCCAT 3141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3093 GGGGTCCGCTTTTCGGGAAGCTCTTCTTTCGCTTATGGGTTTATGGCCATCTCTACCCCT 3152
Qy 3142 TCCTCAAGGCTCTTATGGGAGGCAAAACCGCACACCGAGATGTCTGCTGGGCTG 3201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3153 TCTTGAAGGCTCTTGGGAAGCAAAACCAACCAACCAACCAATCTGTTCTGCTGCTG 3212
Qy 3202 TCCTCTCGCTTCTATCTTCTCTGCTGCTGGGTTCTGTTGTATTCCTATT 3250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3213 TTCTTCTCGCTCCATCTTCTGCTTGGTGGTTCAGATCAATCCCTTT 3261
```

RESULT 4

```
US-08-960-048-1
; Sequence 1, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-1
```

```
Query Match 33.5%; Score 1215.6; DB 3; Length 3328;
Best Local Similarity 69.2%; Pred. No. 0;
Matches 1710; Conservative 0; Mismatches 739; Indels 21; Gaps 3;
```

```
Qy 807 GAGCCTTACTGATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCATTCCTCC 866
Db 617 GAAGATAAACCGGACCGGATGCTTCCAGCCCTCTCGACTATAATTCCAATCCCGAAA 676
Qy 867 TCCAAATAAATCCCTACAGAATGGTCAATTTGTTCTGCGGTTGGTGTGTTTAAGCATCTTC 926
Db 677 AGCAGACTTGCAACCATACCGAACCGGATCATTTATGCGATTGATCATCTTGGTCTTTTC 736
Qy 927 CTGCACTACGCTCTCAAAATCTGTGCGGTAATGTCATACCACCTGTGGCTTTTATCTGTT 986
```


Db 737 TTCCATTATCGAGTAACAAACCCGGTTGACAGTGTCTTTTGGACTGTGGCTCACTTCAGTC 796
QY 987 ATATGTGAGATTGGTTTGGTTTATCTCGGACTGATGATCAGTTCCTCCGAAAGTGGTTTCCA 1046
Db 797 ATATGTGAATCTGGTTTGCATTTTCTGGGTGTGGATCAGTTCCCTAAGTGGTATCCT 856
QY 1047 ATCAACCGGGAGACCTACCTTCATAGACTGGCTTTAAAGGTATGACCGAAGGTGAACCG 1106
Db 857 GTTAACAGGGAACAATACATTCACAGACTATCTGCAAGATATGAAGAGAGGTGAACCT 916
QY 1107 TCTCAGTTGGCTGTCTTGACATATTTGFCAGTACAGTCGACCCCTTGGAAGAGCCACCT 1166
Db 917 GATGAACCTTGCTGCAAGTTGACTTCTTCGTGAGTACAGTGGATCCCATTTGAAAGAGCTCCA 976
QY 1167 ATCGTCACTGCCAACACATGTGTATCCATCTTGCTGTGATATCCCGTGACAAAGTTC 1226
Db 977 TTGATTACTGCCAAATCTGTGTCTTCCATCTTGCCTTGGACTACCCGGTGATTAAGTTC 1036
QY 1227 TCTTGTATGTATCTGATGACGGAGCTTCAATGCTGACTTTTGGACCAATTTGGCTGAGACT 1286
Db 1037 TCTTGTATATATCTGATGATGGTGGGCCATGCTGACATTTGAATCTCTAGTAGAAACA 1096
QY 1287 TCAGAGTTTGTAGGAAATGGGTACCAATTTGTGAAGAGTATGACATTTGAACCCAGAGCT 1346
Db 1097 GCCGACTTTGCAAGAAAGTGGGTTCATTTCTGCAAAAATTTTCCATTTGAAACCCCGGCA 1156
QY 1347 CCGAGTTTACTTTTGCAGAAATTTGATTTACTTACCTGAAAGACAAAGTTCAGCTTCATTT 1406
Db 1157 CTTGAGTTTACTTCTCACAGAAGATTGATTAATCTTGAAGATGAAGTGCAGGCCCTCTTTT 1216
QY 1407 GTTAAAGACCGCGGCCATGAAGAGAGATATGAAGATTTAAATCAGGATAAATGCC 1466
Db 1217 GTAAAGAACGTAGAGCTATGAAGAGATTTGAAGATGACAAATTCGATCATGTCT 1276
QY 1467 CTAGTTTCTAAGSCATTTAAAGTCCCGAGGAGGATGATCATGCAAGATGGCACACA 1526
Db 1277 TTAGTTGCAAGGCTCAGAAAACACCTGATGAAGATGCACAATGCAAGATGGAATCTCT 1336
QY 1527 TGGCAGGAAACATACACAGGATCATCTGGAATGATTCAGGTTTCTTGGTTCAGT 1586
Db 1337 TGGCCAGGAAATAACCGCGGTGATCACCTGGCATGATTCAGGTTTCTTGGATATAGT 1396
QY 1587 GGTGGCTGTGATCTGAGGTATGAGTCCCGGTTTATGTTTATGTCTCGTGAAGAAG 1646
Db 1397 GGTGCTGTGACATCGAAGAAATGAATCTCTCGACTGGTTTACGTCTTAGAGAGAAG 1456
QY 1647 CGTCTGGGTTCCAGCACCAAGAGGCTGGCCATGAATGCCCTTGTTCGTGTCACA 1706
Db 1457 AGACCTGGCTACCAACACCAAAAAGGCTGGTGTGCTGAAAATGCTTTGGTTAGGTTGTCT 1516
QY 1707 GCTGCTCTTACTAATGGAACAATACATGTTGAATCTTGATTTGATGATCAGTACATCAACAAC 1766
Db 1517 GCAGTCTTTACAAATGCTCCCTTCATCTCAATCTTGATTTGTGACCACTATGTTAAACAAT 1576
QY 1767 AGCAAGCTGTCCGAGAAGCTATGCTCTCTTAATGATCCAAACCTAGTCCGCAAGTC 1826
Db 1577 AGCAAGCAGTTAGGAGGCAATGTCTCTTCTGATGGAACCAAGATGTTGAGATGTA 1636
QY 1827 TGTATTGTGCAAGTCCCAAAAGGTTTGTGAGGATTTGATAGGAATGATCGATATGCAAAAC 1886
Db 1637 TGCTATGTGCAAGTTTCTCAAGATTTGATGGCATAGATAGGAGTGTATGATGCCAAT 1696
QY 1887 AGGAACATGTCTTTTGTGATATTAATCTTGAGGGGCTTGAACGGCAATTCAGGAACAGTT 1946
Db 1697 AGGAACACAGTTTCTTGTATGTTAAACATGAAGGTCCTTGATGGAATCCAAAGGGCCAGTT 1756
QY 1947 TATGTGGGAACGTGTTTTCACAGAACAGCTATCTATGTTATGAGCCCCCAAT 2006
Db 1757 TATGTGGGAACAGTTGTGTTTTCATATAGCAAGCACTTTATGGCTATGTTCCACCTTCA 1816
QY 2007 AAGCGGAAGAACGAGGTTTCTTGGCATCATATGTGGGGGCAAGAGAGGCAAGCAAG 2066
Db 1817 ATGCCAAGTTTCCCAAGTCATCTCTCATCTTGTCTGTGTTGCTGCCCGGGCAAGAG 1876

QY 2067 TCAAAGAAAGAGAGCTCAGATAAGAAAAGTCGAACAAGCATGTGCACAGTTCTTGTTCCA 2126
Db 1877 GAACCTAAAGATCCATCAGAGCTTTATAGGATGCAAAACCGGGAAGAACTTATGCTGCC 1936
QY 2127 GTATTCAATCTCGAAGACATAGAGGAGGTGTGAAGGTGCTGGGTTTGATGATGAGAAA 2186
Db 1937 ATCTTTAACTTACCTTAGGAAATTGACAATTATG-----ATGAGTATGAAGA 1981
QY 2187 TCAGTTCTCATGTCTCAATAGAGCTTAGAAGAAGATTTTGGCCAGTTCAGCAGCATTTGTT 2246
Db 1982 TCAATGTTGATCTCTCAAAACAGCTTTGAGAAAACCTTTGGCTTATCTTCAGTCTTCATT 2041
QY 2247 GCCTCCACTCTGATGGAATATGGTGTCTCTCAGTCTCTCAGTCTCCACCTCCAGAACTCTTTTG 2306
Db 2042 GAATCTACACTAATGGAGATGGAGAGTGGCTGAATCTGCCAACCCCTTCCACACTAATC 2101
QY 2307 AAAGAAGCTATCCATGTCTATAAGTTGTGGCTATGAGGACAAAGTCTGAAATGGGAACTGAG 2366
Db 2102 AAGGAAGCAATTCATGTCTCATCAGCTGTGGCTATGAAGAGAGACTGCATGGGGGAAGAG 2161
QY 2367 ATTGGTTGATCTATGGATCTCTCAGAGAATATTTCTTACTGGATTAAGATGCAAGCA 2426
Db 2162 ATTGGATGGATATATGGTTTCACTCAGGATATCTTAAACGGCTTCAAAATGCACTGC 2221
QY 2427 AGAGCTCGGTTCACTATTTGCATGCCAAGCGCCAGCTTTCAGGGATCTGCCCC 2486
Db 2222 CGAGGATGGAGATCGAATTTACTGCAATGCCCTTAAAGCCAGCATTCAGAGATCTGCACCC 2281
QY 2487 ATCAATCTTTCAAGATCTCTGAACCAAGTGTGCGGTGGGCTCTCGGTTCTCTTGAAAT 2546
Db 2282 ATCAATCTGTCTGATCGGTTGCACCAAGTTCCTCGATGGGCTCTTGGATCTGTGAAAT 2341
QY 2547 CTTTTCAGCCGCAATGCCCCCTTATGGTATGCTACGGA---GGGCGCCTCAAGTTCTCTG 2603
Db 2342 TTCTTAAGCAGGCAATTTGCCCTCTATGGTATGGCTTTGGAGGTGGTCTTAAATGGCTT 2401
QY 2604 GAGATTCGCTTACATCAACACCACTTTTACCACTAACCTCTCTCCGCTCTCTAGTC 2663
Db 2402 CAAAGCTAGCATATATAACACCAATGTGTATCTTTTCAATCTCCATCTCATGTGCC 2461
QY 2664 TATTGTATATTCCTGCTATCTGTCTCTCACTGGAAAGTTTCAATGATGCAGAGATTAGC 2723
Db 2462 TATTGTTCACTACAGCAATCTGTCTCTCAAGGAATTTTATCATCAACGCTCTCA 2521
QY 2724 AACTTGGCCAGTATCTGGTTCAATTTGGCTCTCTTTCAATTTTCCCACTCGTATCCCT 2783
Db 2522 AACCTGGCAAGTGTCTCTTCTTGGCTTTTCTTCCATTTATCGTACTGTGTCTC 2581
QY 2784 GAGATCAGGTGAGTGGTGTGGCATTTGACAGTGTGTGAGGAAGTGAACAGTTCTGGGTC 2843
Db 2582 GAGCTCCGATGAGTGGTGTGAGGACTTATGGCGTAAACAGCAGTTTGGGTC 2641
QY 2844 ATTGGAGGTATCTCTGCACATCTGTTTGGCTTTTTCAGGGCTCTTCTGAAGGTGCTTGGC 2903
Db 2642 ATCGTGGGTTTTCAGCCATCTCTTGGCTCTTCCAGGTTTCTTAAAGTGTGCG 2701
QY 2904 GGTATCGACACCACTTCTCATCTGTCACTCAAGGCTTAATGACGAAGAGGCACTTTGCT 2963
Db 2702 GGCATTGACACCACTTTACTGTCTGTCGCAAGCAGCTGATGATGCA---GATTTTGGT 2758
QY 2964 GAGCTCTAGATGTTCAAGTGGAGCGGTTCTCATCTCCCTCCGACGACCAATTTTGATCATT 3023
Db 2759 GAGCTCTAGATGTTGAATGGACTACATTTCTAAATCCCTCCAAACAACTCTCTCATCGTC 2818
QY 3024 AACATGTTGGTGTCTGCTGCGACCTCTTACGCCATCAACAGTGGTTTACCAATCATGG 3083
Db 2819 AACATGTTGGTGTCTGTCGCGGATTTCTCCGATGCCCTCAACAAAGGTACGAGCTGG 2878
QY 3084 GGGCCGCTCTTTGGGAAGCTCTTCTTGGCTCTCTGGGTGAATTTGACATTTATCCCATTC 3143
Db 2879 GGACCACTCTTTGGCAAGTGTCTTCTTCTGCGGTCTCTCTCATCTCTTATCATTCATTC 2938

QY 2367 ATTGGTTGATCTAUGGATCTGTCCAGAGATATTTCTTACTGGAATCAAGATGCACGCA 2426
|||
Db 2162 ATTGGATGATATATUGTTTCAGTCACTGAGGATATCTTAACGGCTTCAAAATGCATGC 2221
|||
QY 2427 AGAGGCTGCGTTTCACTTATTCATGCCCCAGCGCCAGCTTTCAAGGATCTGCCCC 2486
|||
Db 2222 CGAGGATGAGATCGATTTACTGATGCGCTTAAGGCCAGCATTCAAAGGATCTGCACCC 2281
|||
QY 2487 ATCAATCTTTCAGATCGTCTGAACCAAGTCTGCGGTGGCTCTCGTTCTGTGAAAT 2546
|||
Db 2282 ATCAATCTGTCTGATCGGTGACCAAGGTTCTTCGATGGGCTCTTGATCTGTGAAAT 2341
|||
QY 2547 CTTTTCCAGCGGCAATGCCCTTAUGGATGATGCTACGGA---GGGCGCCTCAAGTTCCCTG 2603
|||
Db 2342 TTCCTAAGCAGCAATGCCCTCTATGGTATGGCTTTGGAGTGGTCTGTCTTAATGGCTT 2401
|||
QY 2604 GAGAGATTCGTTTACATCAACCAACCATTTTACCCACTAAGCTCTCTCCCGCTTTAGTC 2663
|||
Db 2402 CAAAGACTAGCATATATAAACACCAATTTCTATCTCTTACATCCCTCCACTCATTCGC 2461
|||
QY 2664 TATTGTATATGCTGCTATCTGTCTGCTCACTGGAAGTTTCATATGCGAGATTAGC 2723
|||
Db 2462 TATTGTCTACTACCAAGCAATCTGTCTCTCACAGGAAATTTATCATCAACAGCTCTCA 2521
|||
QY 2724 AACTGGCCAGTATCTGGTTTCATTGCGCTCTCTCTTCAATTTTCGCCACTGGTATCCTT 2783
|||
Db 2522 AACCTGGCAAGTCTCTCTTTTGGCTTTTCTTCCATATCTGACTGCTGTCTC 2581
|||
QY 2784 GAGATGAGTGGAGTGGTGTGGCATTTGAGATGAGTGGAGGAATGAACAGTTCTGGGTC 2843
|||
Db 2582 GAGCTCCGATGAGTGGTGTGAGCAATGAGGACTTATGGCGTAAAGCAGATTTTGGGTC 2641
|||
QY 2844 ATTGGAGGTATCTCTGCAATCTGTTTGGCTTTTTCAGGCTTTCTGAGGTGCTTGC 2903
|||
Db 2642 ATCGTGGCGTTTCAGCCCATCTCTTTGGCGTCTTCCAGGTTTCTTAAAGATGCTTGG 2701
|||
QY 2904 GGTATCGACACCAACTTCTACTGTCACTCAAGGCTTAATGACGAAGAGCGCACTTTGCT 2963
|||
Db 2702 GGCATTGACACCACTTTACTGTCTGCTGCAAGAGAGCTGATGATGCA---GATTTGGT 2758
|||
QY 2964 GAGCTCTAATGTTCAAGTGGAGCGCTTCTATCCCTCCGACGACCAATTTTGATCAT 3023
|||
Db 2759 GAGCTCTAATGTTGAATGGACTACACTTCTAATCCCTCCAACACACTCCCTCATGTC 2818
|||
QY 3024 ACATGTTGGTCTGCTTCTGCTGCACTCTGACCCATCAACAGTGGTTTACCAATCATGG 3083
|||
Db 2819 AACATGGTGGTCTGCTGCTGCGGATTTCTCGATGCGCTCAACAAAGGGTACGAAGCTGG 2878
|||
QY 3084 GGGCGGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGGTGATTTTCACTTATACCCATTC 3143
|||
Db 2879 GGACCACTCTTTGGCAAGTGTCTTTTCTTCTGGGTATCTCTCATCTTATCCATTC 2938
|||
QY 3144 CTCAGGGTCTTATGGGAGGCAAAACCGCACACCGAGATGTCTATCGTCTGGGCTGTC 3203
|||
Db 2939 CTCAGAGGTCTTATGGGAGGCAAAACAGACACCAACCAATTTTCTCTTGGTCTGCTG 2998
|||
QY 3204 CTCCTCGTCTTCTCTCTCTCTGCTGCTGGGTCTGTTGATCCATCTACTACCGCTCTC 3263
|||
Db 2999 TTGTTGGCTCTGCTCTCTCTCTTCTTGTGGGTTCGATCAACCCGTTTGTGACACCGCC 3058
|||
QY 3264 GCTGGGCCAA 3273
|||
Db 3059 GATAGCACCA 3068
|||

RESULT 6

US-09-221-013A-7
; Sequence 7, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.

; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3828
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (239)..(3490)
US-09-221-013A-7

Query Match 32.4%; Score 1176.6; DB 4; Length 3828;
Best Local Similarity 68.2%; Pred. No. 2.2e-311;
Matches 1667; Conservative 0; Mismatches 764; Indels 12; Gaps 2;
QY 817 TGAATGATGAAGTCCGAGCTCTATCTAGAAAAGTCCCATTTGCTTCTCCCAATAA 876
|||
Db 999 TGATGATGAAGAGAACCACTCTCTCAAGAAAGTACTATTCGTTCAAGCAGATAA 1058
|||
QY 877 ATCCCTACAGAATGGTCTTCTGCGGTTGGTTCTTAAAGCATCTTCTGCACTACC 936
|||
Db 1059 ATCCCTACAGGATGTTAATCTGTGCGCTCGCATCTCTGTCTCTTCTTCAATA 1118
|||
QY 937 GTCCTCAAAATCCTGTGCGTAATGATACCCACTGTGGCTTTTATCTGTATATGAGA 996
|||
Db 1119 GAATTCCTCCATCCAGTCAATGATGATATGATATGATGTTAAACGTGATATATGCGAA 1178
|||
QY 997 TTTGGTTTCTTTTATCTCGTACTGATGATCAGTCTCCGAAGTGGTTTCAATCAACCGGG 1056
|||
Db 1179 TATGTTTTCAGTGCTTGGATCTTGTATCAATTTCCCAATGGTAICTTATAGAACGTG 1238
|||
QY 1057 AGACCTTACCTTGATAGACTGGCTTTAAGGTATGACCGAGAAGGTGAACCGTCTCAGTTGG 1116
|||
Db 1239 AAACATACCTCGATAGACTCTCTCAGGTACGAGAAGGAGGAAACCGTCAGGATTAG 1298
|||
QY 1117 CTGCTGTGACATATTTGTTCAGTACAGTCGACCCCTTGAAGGAGCCACTATGTCACGTG 1176
|||
Db 1299 CACCTGTTGATCTTTTGTGTAGTGGATCCGTTGAAGAGGCCCTTGAATTACAG 1358
|||
QY 1177 CCAACACTGTCTATCTCTTCTGTTGATTTATCCCGTGGACAGGCTCTCTTGTCTATG 1236
|||
Db 1359 CAAACACAGTTCTTTCATCTAGCATGATTAICTGTGGATAAGTTGGTGTGTTATG 1418
|||
QY 1237 TATCTGATGACGAGCTTCAATGCTGCTTTTGAAGCATTTGGCTGAGACTTCAGAGTTTG 1296
|||
Db 1419 TATCAAAACAATGGTGAGCTATGCTTACATTTTGAAGCTCTCTCTGATCAGCTGATTTG 1478
|||
QY 1297 CTAGGAAATGGGTACCATTTTGTGAAGAGTATGACATGAAACCGAGAGCTCCGAGTTT 1356
|||
Db 1479 CTACAAATGGGTTCCTTTTCTGAAGAGTTTAAATATCGAGCCACGAGCTCCTGAGTGT 1538
|||
QY 1357 ACTTTTGGCAGAAAATGATTAACCTGAAAGACAAAGTCCAGCTTCATTTTAAAGACC 1416
|||
Db 1539 ATTTTCTCAGAAGATGGATTACCTGAAGAACAAAGTTCTCTGCTTTTGTGACGGGAAC 1598
|||
QY 1417 GCCGGGCCATGAAGAGAGAAATATGAAGATTTTAAATCAGGATAAATGCCCTAGTTTCTA 1476
|||
Db 1599 GTGCTGCTATGAAGAGATTTATGAAGATTTAAAGTGAAGATAAATGACCTGTTGCTA 1658
|||
QY 1477 AGGCATTTGAAAGTCCCGAGGAGGATGATCATGCAAGATGGCACACCATGGCCAGGAA 1536
|||
Db 1659 CTGCACAGAAAGTGGCTGAGGAAAGTGGACTATGCAAGATGGAATCCTTGGCCTGGAA 1718
|||
QY 1537 ACAATACAGGAGATCATCTCTGGAATGATTCAGTTTCTTCTGTCACAGTGGTGGCCTTG 1596
|||

Db 1719 ACAAAGTCCGTGACCATCTCGAATGATCAGGTGTTCTTGGGTCATAGTGGAGTTCGTG 1778
QY ATACTAGGGTAATAGAGTCCCGCTTTAGTTTATGTGTCGCGTGAAGAGCGTCTCGGT 1656
Db 1779 ATACGGATGTAATAGAGTTACACAGTCTAGTGTATGTTCTCGTGAGAACGCGCTGGAT 1838
QY TCCAGACCAACAAGAGCGTGGCCATGAATGCCCTTGTGTCGTTGCTCAGCTGTCCTTA 1716
Db 1839 TTGATACCAACAAGAGAGCTGGAGCTATGAATTCCTTGATCCGAGTCTCTGCTGTTCTAT 1898
QY CTAATGACCAATACATGTTGAATCTTGATTTGATCACTACATCAACAAACAGCAAGGCTG 1776
Db 1899 CAACGCTCCTTACCTTCTTAATCTCGATCTGATCACTACATCAACAAACAGCAAGCAA 1958
QY TCCGAGAGCTATGCTTCTTAATGATCAAAACCTAGTTCGGAAGTCTGTTATGTGC 1836
Db 1959 TTAGAGAACTATGTTGTTTATGATGACCCGCAATCGGGAAGAAAGTTTGTATGTTTC 2018
QY AGTTCCCAAAAGTTTATGAGGATTCATAGGAATGATCGATATGCAAAACAGGAACACTG 1896
Db 2019 AGTTCCGACAGATTTATGAGGATTTATAGACATGATAGTACTCAAAACGTAACGTTG 2078
QY TCTTTTGTATTAATCTAGGGGCTTGAAGGCTTCAAGGACCAAGTTTATGTGGAA 1956
Db 2079 TGTTCTTTGATATTAACATGAAGAGTCTTGATGGATACAGGACCGATATATGTCGGGA 2138
QY CTGTTGTGTTTCAACAGAACAGCTATCTATGTTATGATGACCCCGCAATTAAGCGGAAGA 2016
Db 2139 CAGTTGTGTTTATGAGAAACAGGCTCTTTATGTTTGTATGACCAAAAGAAAGAAAC 2198
QY 2017 AGCCAG-----GTTCTTGGCAATCACTATGTTGGGGCAAGAAAGGCAAGCAAGT 2067
Db 2199 CACAGGCAAAACCTGTAACCTGTTGGCTTAATGTTGTTGTTGGGTTGAA 2258
QY 2068 CAAGAAGAGGAGCTCAGATAAGAAAAAGTCAAGCAATGTCGACAGTTCCTGTTCCAG 2127
Db 2259 AGAAGAGTAAACGAAAGCCACAGATAAGAAACATAACATAAGAGACTTCAAAAGCAGA 2318
QY 2128 TATTCATCTCGAAGCATAGAGAGGTTGTTGAAGTGTGCGGTTGATGATGAGAAT 2187
Db 2319 TTATGCGCTAGAGAATGTGCAGGAAGTGTATCTGCTCCAGTGTCAAAATGTTGAGAAGA 2378
QY 2188 CAGTTCTCATCTCTCAATAGCTTAGAGAAGAGATTGGCCAGTCAGCAGCATTTGTTG 2247
Db 2379 GATCTGAAGCAACATTAAGATTGAGAGAGAGTTTGGACATCTCCGTTTTCGTTG 2438
QY 2248 CTTCCACTCTGATGGAATATGTTGTTCTCTAGTCTCTCACTCCAGAAATCTCTTTGA 2307
Db 2439 CTTCTGCTTCTACAGAACGTTGGAGTTCCCGCTAACGCAAGCCCGCATGTTTGTAA 2498
QY 2308 AAGAGCTATCCATGTCATAGTTTGGCTATGAGCAAGTCTGAATGGGAACTGAGA 2367
Db 2499 GAGAAGCAATCAAGTTATAGTGGGGTTACCAAGATAAAACCGAATGGGGAAGAGA 2558
QY 2368 TTGTTGGATCTATGATCTCTACAGAAATATTTCTACTGGATTCAAGATGACGCA 2427
Db 2559 TCGGTGATTTATGATCGTGTACTGAAGATATCTGACGGTTTCAAGATGATGCC 2618
QY 2428 GAGGTGCGTGTAGTCTATTTGATGCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCCA 2487
Db 2619 ATGGATGAGATCTGTGTACTGTATGCTTAAGCGTGCAGCTTTTAAAGGATCTGTCCTA 2678
QY 2488 TCAATCTTTCAGATCTGTGAACCAAGTCTGCGTGGCTCTCGTCTCTGTAATTC 2547
Db 2679 TTAATCTTCAGATCTGTTTCAATCAAGTTTACGTTGGGCTCTTGGCTCTGAGAGATT 2738
QY 2548 TTTTCAGCGGCAATGCCCCCTTATGTTATGCTACGAGGCGGCTCAAGTTCTCTGGAGA 2607
Db 2739 TCTTCAGCAGACATTTGTCGATATGTTATGTTATGTTGTTGTTTAAATGTTGGAGA 2798
QY 2608 GATTGCGTTTACATCAACACCACTATTAACCACTAACCTCTCTCCGCTTCTAGTCTATT 2667

Db 2799 GATTCTCTTACATCAACTCTGTCGTCTATCCTTGGACTTCACATTCCATTCGTCGTATT 2858
QY 2668 GTATATTGCGTCTATCTGTCGTCTCACTGGAAGTTTCATCATCCAGAGATTAGCAACT 2727
Db 2859 GTTCTCTCCCGCGGTTTGTACTACAGAAAAATTCATGCTCCCTGAGATAAGCAACT 2918
QY 2728 TGGCCAGTATCTGTTTCAATTCGCTCTTCTTCAATTTTTCGACACTGGTATCTCTGAGA 2787
Db 2919 ACGCAGGTATCTTCTCATGCTCATATCCATAGCAGTAACTGGAACTCTCGAAA 2978
QY 2788 TGAGTGGAGTGTGTTGGCATTGACAGTGGTGGAGAAATGAACAGTCTGGGTCAATTG 2847
Db 2979 TCGAATGGGAGGTGTCGGAATCGATGATTGGTGAGAAACGAGCAGTTTGGGTAAATCG 3038
QY 2848 GAGTATCTCTGACATCTGTTGCGCTCTTTCAGGGTCTTCTCAAGGTCTTCAAGGTCTTCCCGGTA 2907
Db 3039 GAGGGGCTCTCTGCACTATTTGCTCTGTTTCAAGTCTTCTCAAGTCTTAGCCGAG 3098
QY 2908 TCGACACCAACTTCACTGTCACTCAAGGCTTAATGACGAAGAGGCGACTTTTCTGAGC 2967
Db 3099 TTAACACGAATTTTCACAGTCACTTCAAAAGC--AGCAGACGATGAGCTTCTCTGAGC 3155
QY 2968 TCTACATGTTCAAGTGAACGCGCTTCTATCCCTCCGACGACCATTTTGTATCAATTAACA 3027
Db 3156 TTTACATCTTCAAGTGAACAACTTTGTTGATTTCTCCGACAACTTCTGTATCAATTAACA 3215
QY 3028 TGGTTGTTGCTGTTGCTGGCACCCTTACGCCATCAACAGTGGTTACCAATCATGGGGC 3087
Db 3216 TCAATGAGTATTTGTCGGGCTTCTGATGCCATTAGCAATGGGTATGACTCATGGGAC 3275
QY 3088 CGCTCTTTGGGAAGCTCTCTTTTGGCTTCTGGGTGATTGTTTCACTATACCCATTCTCTCA 3147
Db 3276 CTCTTTGGGAGACTTTTCTCTGCTCTTTGGGTCAATTTTCAATTAATACCCATTCTCTCA 3335
QY 3148 AGGTCTTATGGGAGGCAAAACCGCACCGAGATGTCATCGTCTGGGCTGCTCTCC 3207
Db 3336 AGGGAATGCTTGGGAAGCAACAAAATGCTACGATTTATGTGCTGTGTTCTTCTTC 3395
QY 3208 TCGCTTCTATCTCTCTCTGCTGCTGCTGCTGTTGATCCATT 3250
Db 3396 TAGCTTCGATCTTGACACTCTTGTGGGTGAGAAATAACCCGTT 3438

RESULT 7

US-09-221-013A-1
; Sequence 1, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2448
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1887)
US-09-221-013A-1

Query Match 27.1%; Score 984.2; DB 4; Length 2448;
Best Local Similarity 72.1%; Pred. No. 7,1e-259;

SEQ ID NO 13
LENGTH: 1741
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1741)
US-09-221-013A-13

Query Match 15.7%; Score 568.8; DB 4; Length 1741;
Best Local Similarity 62.0%; Pred. No. 2.7e-145;
Matches 1031; Conservative 0; Mismatches 547; Indels 84; Gaps 5;

Qy 53 CCGCGCATGACGCGACGCGGACGCGCTGAGCTCCGGAGGACACCGGGCCGGGCGT 112
Db 160 CCGCCCCGACGCGGACGCGCCACCGCGGCTTAAGCGAGGAGAGTGTGAATGTGAGT 219
Qy 113 GTGCCAGATCTGCGCCGACGCGCTGGGCACACAGCTTGACGCGGACGCTTTCACCGCTG 172
Db 220 CTGCCAGATTTGTGGGACACTGTTGGGCTCTCGGCCACCGCGGACGCTTTTGTGCTG 279
Qy 173 CGAGCTGTGCGCTTCCGGTCTGCGGCTCTGAGTCCAGGACGCGGACGCGGAGGCGAC 232
Db 280 CAATGAGTGGCGCTTCCGGTCTGCGGCTTGTCTAGGATACGAAACGAGGAGGAA 339
Qy 233 CCAGGCTGCTCCAGTGCAGACCAAGTACAGCGCCACAGAGGAGGCGGCGATCCG 292
Db 340 CCAGTGTGCCCCAGTGCAGACCTAGATACAGAGGCGCAAGGTTGCCCTAGAGTTCA 399
Qy 293 CGGGAGGAGGCGGACGACACTGATGCGGATGATGTAGTGAATTCACCTACCTGTCATC 352
Db 400 GGGCGATGAGGAGAGAGATGTTGATGACCTGACCAATGAATTCATTA----- 450
Qy 353 TGGCACTGAGGACCAAGAGAGAGATGCTGAAGGATGCGAGCTGGCGCATGAACAC 412
Db 451 -----TAAGCATGG 459
Qy 413 CGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATGGCGCTTCCAA 472
Db 460 CAATGGGCAAGTCCAGTGGCGAGATACAGAGACAGGGGAGAGATGTGACCTGTCTTC 519
Qy 473 GTATGACAGTGGAGAGATCCCTAGGGGATAGCTCCCTTCAGTCAACCAAGCGGAGATGC 532
Db 520 ATCTCTCGCCACGAACAACATCGAATCCCGTCTGACAAAGTGGCA---ACAGATCTC 576
Qy 533 AGGAGAAATCCCTGAGGCTTCCGCTGATCATCATGATGTCCCTACGGGAGCATCAG 592
Db 577 AGGAGAGATCCCTGATGCTTCCCGCATCGCCATTTCTATCCGCGAGCGGAAACATCAAGCTA 636
Qy 593 CAGAGCTGCTCGGTTTCCCTATGTAATCAATCCAAATCCGTCAAGGGAGTTC---TC 649
Db 637 TGTGATCCAGTGTTCAGTTCCTGTGAGGATTTGGACCCCTCAAGGACTTGAATTC 696
Qy 650 CGGCAGTATTGGGAATGTTGCTGGAAAGAGAGAGTGTGATGGCTGGAAATGAAGCAGA 709
Db 697 CTATGGGATTAACAGTGTGACTGCGCAAGAAAGAGTTCAGCTGGAGGAAACAGCAGA 756
Qy 710 CAAGGTCGATTCCTGACTAATGGGACAGCAATGCTCCCTCTGAGGTTCGGGACG 769
Db 757 CAAAAATATGATGCAAGGTAGCTAAT-----AAATATCCAGAGGCAAGAGGGGAGA 807
Qy 770 TACTGACATGATCATCTACTGATACAAATGAGAGAGCGCTTTTACTGAATGATGAAC 829
Db 808 CATGA-----AGGACTGGTTCAAATGGTGAAGATATCCAAATGTTGATGATGC 858
Qy 830 TGCAGAGCTCTATCTAGAAAGTCCCATGCTTCTCCCAAAATATAATCCCTACAGAT 889
Db 859 AGCTCTACTCTGAGCGGATAGTGCCTATCCCTTCAACACAGCTCAACCTTTACCGAT 918
Qy 890 GGTCAATGTTCTGGGTTGGTGTCTAAGCATCTTCTGCACTACCGTCTCACAAATCC 949
Db 919 TGTTATCATCTCCGCTTATCATCTCTGATGTTCTTCTTCCAAATATCGGTGCTCATCTCC 978

Qy 950 TGTGCGTAATGCAATACCACTGTGGCTTTTATCTGTATATGTGAGATTTGGTTGCTTT 1009
Db 979 AGTGGCGGATGCTTATGATTTGGCTAGTATCTGTATCTGTGAAATTTGGTTGCCCTT 1038
Qy 1010 ATCTGTGATATCTGATCATGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGA 1069
Db 1039 ATCTGTGCTCTAGATCAATTTCCCAAAGTGGTACCCGATAAACCGTGAACATACCTTGA 1098
Qy 1070 TAGACTGCTTTAAGGTATGACGAGAGGTGAACCGCTCTCAGTTGGCTGTGTGTGACAT 1129
Db 1099 CAGGCTTGCAATGATATGATGAGGAGGAGGACCAATCAAGCTTCTCTCCATGTATGT 1158
Qy 1130 ATTGTGCACTACGTCGACCCCTTTGAAGGAGCCACTATCTGCTACTGCCAACACTGTGCT 1189
Db 1159 CTTTGTGAGTGGTGGATCCACTTAAAGGAACCTCTCTGATCAACAGCAAACTGTGTTT 1218
Qy 1190 ATCCATTTCTGCTGTTGATATCCGCTGGGACAGGCTCTCTGCTATCTATCTGTATGAGCG 1249
Db 1219 GTCCATTTCTGCTGTTGATTAACCTGTGACAAAGTGTCTGCTATGTTCTGACGATGG 1278
Qy 1250 AGCTTCAATGCTGACTTTTGAACGCAATGGCTGACACTTCAGAGTTTGTAGGAAATGGGT 1309
Db 1279 TTGAGCTATGTTAATCTTTGAGGCTCTGTGAGAACTGCGAATTTGCTAGGAAAGTGGT 1338
Qy 1310 ACCATTTGTGAGAGAGTATGACATTGAACCCAGAGCTCCGAGTTCCTTTTCCAGAA 1369
Db 1339 TCCGTTTTGCAAGAGCAACAATTTGAACACAGAGCTCCAGAGTTTACTTTTGTCTCAAA 1398
Qy 1370 AATTGATTAATCTGAAAGCAAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGCGCATGAA 1429
Db 1399 AATAGATTAATCTGAAAGCAAAATCCAACTTCTTTGTTAAAGAAAGGCGGCGCAATGAA 1458
Qy 1430 GAGAGATATCAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCAATGAAAGT 1489
Db 1459 GAGAGATATGAGAGATTTCAAGTACCGATCAATGCTCTTGTGCGAAGGCAACAAAGT 1518
Qy 1490 CCCCAGGAGAGGATGATGCAAGATGCGACACCAATGCGGAGGAGGAGGAGGAGGAGGAG 1549
Db 1519 ACCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1578
Qy 1550 TCATCTGGAATGATCAGGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1609
Db 1579 TCACCTGCGCATGATTCAGGTGTTCTTGGGGGACAGTGGTGGGCTTGGACCTGATGCTG 1638
Qy 1610 TGAGTCTCCCGCTTTAGTTTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1669
Db 1639 CGAGTTCACCGCTTGTCTAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1698
Qy 1670 GAAGGCTGGTCCCATGAATGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
Db 1699 GAAGGCTGGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740

RESULT 9

US-09-221-013A-3
; Sequence 3, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

6817 TTGGATCTATCGAGATCTTTCTAGCAGACATTTCTCTATCTGATGTTACCATGGA 6876
2589 CGCCTCAAGTTCTGGAGAGATTCGCTTACATCAACACCACTTTTACCCATCACTCT 2648
6877 AGGTTGAGACTTTTGGAGAGATCGCTTATATCAACACCATCTCTATTTACATCC 6936
2649 CTCGCGCTTCTAGTCTATTTGCTATATTTGCTCTCTATCTCTGCTCACTGGAAGTTCTATC 2708
6937 ATCCCTCTTATTTGGGTATTTATTTCTCCCGCTTTTGTCTCATCAACGACAGATTCATC 6996
2709 ATGCCAG----- 2715
6997 ATACCCGAGGTTGTAAACCTGACACACTGCTGTAATTTACTATTTGAATCCCATTTTGTGA 7056
2716 -----AGATTAGCAACTTGGCCAGTATCTGTTTCAT 2746
7057 ATGCATTTTGTTCATCATCATCTGTTGACAGATAAGCAACTACGCGAGTATTTGTTTCAT 7116
2747 TGGCTCTTCTTTCAATTTTGGCACTGCTATCTTGGATGAGTGGAGTGGTGGTGG 2806
7117 TCTACTCTTCTATCTCAATTTGCTGTGACTGGAATCTCTGGAGCTGAGATGGAGCGTGGAG 7176
2807 CATTGACAGTGGTGGAGGAAATGAACAGTCTCTGGGTCAATTTGGAGGTATCTCTGCACATCT 2866
7177 CATTGAGATTTGGTGGAGGACGAGCAGTTCTGGGTCAATTTGGTGGACATCCGCCATCT 7236
2867 GTTTGCCGCTTTTCAAGGCTTTCTGAAGTCTTCCCGGTATCGCCGATTCGACCACTTCACTGT 2926
7237 TTTTGTCTCTTCCCAAGCTCTACTTAAGGTTCTTGTCTGATTCGACACCACTTCCCGT 7296
2927 CACCTCAAGGTAATGACGAGAGAGGAGCTTTTGTCTGAGCTCTACATGTTCAAGTGGAC 2986
7297 TACATCTAAGCCACAGAGATGGGATTTTGGAGAACTCTACATCTTCAAAATGGAC 7356
2987 GACGCTCTCTATCCCTCCGACGACCAATTTTGAATCAATTAACATGTTGGTGTCTGTTGCTGG 3046
7357 AGCTCTTCTATTCACCAACACCGCTCTACTTGTGAACCTCATAGCAATTTGGCTGG 7416
3047 CACCTCTACGCCATCAACAGTGGTTTCAATATCATGAGGGGCGCTCTTTGGGAAGCTCTT 3106
7417 TGCTCTTATGCTGTAACAGTGGTACCACTGCTGAGTGGGCTCGCTTTTGGGAAGCTCTT 7476
3107 CTTTGGCTCTGGGTGATTTGCTATATACCATTTCTCAAGGCTCTATCGGAGGCA 3166
7477 CTTGCGCTTATGGGTATTTGCCCATCTTACCATTTCTTGAAGGCTCTTTGGGAAGCA 7536
3167 AAACCGCACACCGAGATTTGCTATGCTCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 3226
7537 AAACCGAACACCAACCATGCTCATTTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7596
3227 GCTGTGGGTCTGTTGATTCATT 3250
7597 GCTTTGGGTGAGATCAATCCCTT 7620

RESULT 10
US-09-221-013A-4
; Sequence 4, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27

; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(5009)
; OTHER INFORMATION: N is A, T, G or C.
US-09-221-013A-4

Query Match 6.1%; Score 219.6; DB 4; Length 5009;
Best Local Similarity 60.4%; Pred. No. 1.7e-49;
Matches 437; Conservative 0; Mismatches 199; Indels 88; Gaps 1;

QY 813 TTACTGATGATGAACCTCGCCAGCCTCTATCTAGAAAAGTCCCATGCTCTCTCCAAA 872
DB 2416 TTCAGGATGGATGAAGGAAGAACACCTCTCTCAAGAAAGCTACCTATTTCTGTTCAAGCAGA 2475
QY 873 ATAAATCCCTACAGAATGGTCAATTTCTGCGGTTGGTTGTTCTAAAGCATCTTCTCGAC 932
DB 2476 ATAAATCCCTACAGAATGGTCAATTTCTGCGGTTGGTTGTTCTAAAGCATCTTCTCGAC 2535
QY 933 TACCGTCTCAAAAATCCCTGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 992
DB 2536 TATAGATTTCTCCATCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2595
QY 993 GAGATTTGGTTGCTTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1052
DB 2596 GAGATATGGTTGCGAGTCTTGGATTTCTGATCAATTTCCCAAAATGGTATCTTATAGAA 2655
QY 1053 CGGAGAGCTACTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 2656 CGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2715
QY 1081 -----TAAG 1084
DB 2716 TGAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2775
QY 1085 GTATGACGAGAGTGAACCGTCTCAGTTGGCTGCTCTGATGATGATGATGATGATGATGATGATGAT 1144
DB 2776 GTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2835
QY 1145 CGACCCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1204
DB 2836 GGATCCGTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2895
QY 1205 TGATTTATCCGTTGGAAGAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
DB 2896 TGATTTATCCGTTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2955
QY 1265 TTTTGAAGCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1324
DB 2956 ATTTGAAGCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3015
QY 1325 GTATGAT 1384
DB 3016 GTTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3075
QY 1385 AGACAAAGTCCAGGCTTCAATTTGTTAAAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1444
DB 3076 GAACAAAGTTCATCTCTGCTTTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3135
QY 1445 ATTT 1448
DB 3136 TTTT 3139

RESULT 11
US-09-313-294A-6485
; Sequence 6485, Application US/09313294A
; Patent No. 6476212

```
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO: 6485
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351882H1
; US-09-313-294A-6485

Query Match          5.7%; Score 206.4; DB 4; Length 281;
Best Local Similarity 83.6%; Pred. No. 1.3e-46;
Matches 234; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 2618 CATCAACACCAACATTACCCACTTAACCTCTCTCCCGCTTCTAGTCTATTGTATATTGCC 2677
Db 1 CATCAACACCAACCATCTACCGCTCAGTCCCTCCCGCTCTCTCATTTACTGTATCTCTGCC 60

QY 2678 TGTATCTGTCTGCTCACTGGAAGTTCATCATGCCAGAGATTAGCAACTTGGCCAGTAT 2737
Db 61 TGGCATCTGCTGCTCACTGGAAGTTCATCATGCCAGAGATTAGCAACTTGGCTAGTAT 120

QY 2738 CTGGTTCAATGGCTCTCTCTTCAATTTTGGCACTGGTATCTTGAGATGAGGTGGAG 2797
Db 121 CTGGTTCAATCTCTCTCTTCAATCTCGATCTTGGCACGGGTATCTGGAGATGAGGTGGAG 180

QY 2798 TGTGTTGGCAATGACGAGTGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTC 2857
Db 181 CGSGCTGGGATCGACGAGTGTGGAGGAACGAGCAGTCTGGGTATCGAGGCAATCTC 240

QY 2858 TGCACATCTGTTGGCGCTCTTTCAGGTCTTCTGAAGGTG 2897
Db 241 CGCCCACTCTGCGCGTCTTCCAGGCGCTCTCTCAAGGTG 280

RESULT 12
US-09-313-294A-4753
; Sequence 4753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO: 4753
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348982H1
; NAME/KEY: unsure
; LOCATION: 64, 189
; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-4753

Query Match          5.6%; Score 204; DB 4; Length 300;
Best Local Similarity 82.2%; Pred. No. 6.2e-46;
Matches 245; Conservative 0; Mismatches 52; Indels 1; Gaps 1;
```

```
QY 2716 AGATTAGCAACTTGGCGAGTATCTGGTTCAATTCGGCTCTCTCTTCAATTTGCGCACTG 2775
Db 3 AGATCAGCAACTTGGCGAGCATCTGGTTCAATCTCCCTCTTCAATCTCGATCTTGGCCACGG 62

QY 2776 GTATCCTTGAGATGAGTGGAGTGTGTTGGCAATTCAGAGTGTGGAGGAATGAACAGT 2835
Db 63 GNATCCTTGAGATGAGTGGAGCGGGGTGGGCAATTCAGCAGTGTGGAGGAACGAGCAGT 122

QY 2836 TCTGGGTCATTTGAGGTATCTCTGCACATCTGTTTGCCTGCTTTCAGGGTCTTCTGAAG 2895
Db 123 TCTGGGTGATCGGGGGCATCTCCGCGCACCTCTTGGCCGTGTTCCAGGGCTGTGCTCAAGG 182

QY 2896 TGCTTGGCGGTATCGACCAACTTCACTGTCTCACTCAAGGTATGAGGAAGGCG 2955
Db 183 TGCTGNGCGCATCGACCAACTTCACTGTCTCACTCAAGGTATGAGGAAGGCG 242

QY 2956 ACTTGTGCTGAGCTTACATGTTCAAGTGGAGCAGCTTCTCATCCCTCCGACGACCAT 3013
Db 243 ACTTCGGGA-CTGTACATGTTCAAGTGGAGCAGCTTCTTATCCCGCCACCAACCAT 299

RESULT 13
US-09-313-294A-5519
; Sequence 5519, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO: 5519
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350285H1
; NAME/KEY: unsure
; LOCATION: 6, 40, 54, 65, 67, 200, 221, 230, 238, 261
; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-5519
```

```
Query Match          5.5%; Score 198.2; DB 4; Length 291;
Best Local Similarity 81.6%; Pred. No. 2.3e-44;
Matches 235; Conservative 0; Mismatches 52; Indels 1; Gaps 1;

QY 2428 GAGGCTGGGGTTTCAGTCTATTGTCATGCCCAAGCGCCAGCTTTCAGAGGATCTGCCCCCA 2487
Db 4 GANGCTGGCGATCAATCTACTGTCATGCCCTAAGCGACNAGCTTTCAGAGGANCCTGCTCCTA 63

QY 2488 T-CAATCTTTCAGATCGTCTGAAACAGTGTGCGGTGGGCTCTCGGTGTTGAAAT 2546
Db 64 TNCNACCTTTCGATCGTTTGAATCAAGTGTCTCGGTGGGCTCTTGGTTCATTTGAAAT 123

QY 2547 CTTTTCAGCGCGCATTCGCCCTTATGCTATGCTACGGAGGGCGCTCAAGTTCCTGGAG 2606
Db 124 CTTTTCAGCGCGCATTCGCCCTTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 183

QY 2607 AGATTGCTTATACCAACCAACCACTTATACCACTTACCTCTCTCCGCTTCTAGTCTAT 2666
Db 184 AGATTGCTTATACCAACCAACCACTTATACCACTTATACCACTTATACCACTTATACCACT 243

QY 2667 TGTATATTGCTGCTATCTGTCTGCTCACTGGAAGTTCATCATGCCA 2714
Db 244 TGCATATTGCGCAGCAGTGTCTTCTCACTGGAAGTTCATCATGCCA 291
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 05:27:52 ; Search time 1528 Seconds

(without alignments)
11656.672 Million cell updates/sec

Title: US-09-900-237A-29

Perfect score: 3626

Sequence: 1 gcacagagaaccccgctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3228839 seqs, 2456066551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3626	100.0	3626	9	US-09-900-237-29
2	2694.4	74.3	3962	17	US-10-437-963-72402
3	2595	71.6	3799	13	US-10-627-132-21
4	2595	71.6	3799	15	US-10-267-459-5
5	2595	71.6	3799	15	US-10-209-059-21
6	2572.6	70.9	3746	15	US-10-160-719-29
7	2572.6	70.9	3746	15	US-10-160-719-49
8	2572.6	70.9	3773	13	US-10-627-132-9
9	2572.6	70.9	3773	15	US-10-209-059-9
10	2572.6	70.9	3773	15	US-10-160-719-5
11	2326.2	64.2	3704	13	US-10-627-132-13
12	2326.2	64.2	3704	15	US-10-209-059-13
13	2326.2	64.2	3704	15	US-10-160-719-57
14	2319.8	64.0	3534	13	US-10-425-114-2553

c	15	2316.4	63.9	4264	17	US-10-437-963-39762	Sequence 39762, A
	16	1681.6	46.4	1733	9	US-09-900-237-19	Sequence 19, Appl
c	17	1673.6	46.2	3614	15	US-10-229-193-9	Sequence 9, Appl
	18	1431.4	39.5	3191	13	US-10-424-599-95990	Sequence 95990, A
	19	1408.4	38.8	3725	13	US-10-627-132-45	Sequence 45, Appl
	20	1408.4	38.8	3725	15	US-10-209-059-45	Sequence 45, Appl
	21	1408.4	38.8	3725	15	US-10-160-719-13	Sequence 13, Appl
	22	1408.4	38.8	3725	15	US-10-160-719-21	Sequence 21, Appl
	23	1408.4	38.8	3725	15	US-10-160-719-41	Sequence 41, Appl
	24	1383.6	38.2	3851	16	US-10-393-840-69	Sequence 69, Appl
	25	1378.2	38.0	3723	13	US-10-424-599-106896	Sequence 106896,
	26	1374	37.9	3753	15	US-10-160-719-33	Sequence 33, Appl
	27	1374	37.9	3753	15	US-10-160-719-53	Sequence 53, Appl
	28	1374	37.9	3780	13	US-10-627-132-1	Sequence 1, Appl
	29	1374	37.9	3780	15	US-10-209-059-1	Sequence 1, Appl
	30	1374	37.9	3780	15	US-10-160-719-9	Sequence 9, Appl
	31	1364.4	37.6	3563	9	US-09-900-237-25	Sequence 25, Appl
	32	1343.4	37.0	3603	15	US-10-229-193-5	Sequence 5, Appl
	33	1341.8	37.0	3673	15	US-10-229-193-11	Sequence 11, Appl
	34	1340.2	37.0	3786	9	US-09-900-237-7	Sequence 7, Appl
	35	1340.2	37.0	3813	13	US-10-627-132-17	Sequence 17, Appl
	36	1340.2	37.0	3813	15	US-10-209-059-17	Sequence 17, Appl
	37	1340.2	37.0	3813	15	US-10-160-719-25	Sequence 25, Appl
	38	1340.2	37.0	3813	15	US-10-160-719-45	Sequence 45, Appl
	39	1331.4	36.7	2830	13	US-10-627-132-5	Sequence 5, Appl
	40	1331.4	36.7	2830	15	US-10-267-459-1	Sequence 1, Appl
	41	1331.4	36.7	2830	15	US-10-209-059-5	Sequence 5, Appl
	42	1320.4	36.4	3776	9	US-09-900-237-3	Sequence 3, Appl
	43	1317.2	36.3	3568	13	US-10-627-132-41	Sequence 41, Appl
	44	1317.2	36.3	3568	15	US-10-209-059-41	Sequence 41, Appl
	45	1317.2	36.3	3568	15	US-10-160-719-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-900-237-29
; Sequence 29, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 3626
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-900-237-29

Query Match 100.0%; Score 3626; DB 9; Length 3626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACAGGAAACCCCGCTCCAGCTCTGTGTCGGTGGATCGCTCTGCCCGCA 60

Db 1 GCACAGGAAACCCCGCTCCAGCTCTGTGTCGGTGGATCGCTCTGCCCGCA 60

Qy 61 TGGACGGCGACGGGACCGCCCTTGAAGTCGGGAGGACGGGCGGGACGTGTGCCAGA 120

Db 61 TGGACGGCGACGGGACCGCCCTTGAAGTCGGGAGGACGGGCGGGACGTGTGCCAGA 120

Qy 121 TCTGGCGCGACGGGCGCTGGGACCGACCTGTTGGACGGGACGTCTTACCCGCTCGACGCT 180

QY	2341	AGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAGATA	2400
Db	2341	AGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAGATA	2400
QY	2401	TTCTTACTGGATTCAAGATGCACGCAAGAGGCTGGCGTTTCAGTCTATTGCATGCCCAAGC	2460
Db	2401	TTCTTACTGGATTCAAGATGCACGCAAGAGGCTGGCGTTTCAGTCTATTGCATGCCCAAGC	2460
QY	2461	GCCAGCTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAAACCAAGTCTGC	2520
Db	2461	GCCAGCTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAAACCAAGTCTGC	2520
QY	2521	GTGGGCTCTCGTCTCTGTGTGAAATCTTTTCAGCGGCATTCGCCCTTATGGTATGGCT	2580
Db	2521	GTGGGCTCTCGTCTCTGTGTGAAATCTTTTCAGCGGCATTCGCCCTTATGGTATGGCT	2580
QY	2581	ACGGAGGGCGCTCAAGTTCTGGAGAGATTGCTTACATCAACACCAATTTACCCAC	2640
Db	2581	ACGGAGGGCGCTCAAGTTCTGGAGAGATTGCTTACATCAACACCAATTTACCCAC	2640
QY	2641	TAACTCTCTCGGCTCTAGTCTATTGTATATTGCTGCTGATCTGTCTGCTCACTGGAA	2700
Db	2641	TAACTCTCTCGGCTCTAGTCTATTGTATATTGCTGCTGATCTGTCTGCTCACTGGAA	2700
QY	2701	AGTTCAATCATGCGAGATTAGCACTTCGCCAGTATCTGTTTCAATGGCTCTTCTCTTT	2760
Db	2701	AGTTCAATCATGCGAGATTAGCACTTCGCCAGTATCTGTTTCAATGGCTCTTCTCTTT	2760
QY	2761	CAATTTTGGCACTGGTATCTTGTGATGAGGTGGTGTGGCATTTGACGAGTGT	2820
Db	2761	CAATTTTGGCACTGGTATCTTGTGATGAGGTGGTGTGGCATTTGACGAGTGT	2820
QY	2821	GGAGGAATGAAAGTCTGGGTCAITGGAGGTATCTCTGCATCTGTTTGGCGTCTTTC	2880
Db	2821	GGAGGAATGAAAGTCTGGGTCAITGGAGGTATCTCTGCATCTGTTTGGCGTCTTTC	2880
QY	2881	AGGGTCTCTGAAGTGTCTTGGCGGTATCGACCAACTTCATCTGCCTCAAAAGGCTA	2940
Db	2881	AGGGTCTCTGAAGTGTCTTGGCGGTATCGACCAACTTCATCTGCCTCAAAAGGCTA	2940
QY	2941	ATGACGAAGAGGCGACTTTTGTGAGTCTACATGTTCAAGTGGAGCGCTTCTCATCC	3000
Db	2941	ATGACGAAGAGGCGACTTTTGTGAGTCTACATGTTCAAGTGGAGCGCTTCTCATCC	3000
QY	3001	CTCCAGACCAATTTGATCATTAACTAGTGGTGTCTGCTGCTGCACTCTCTACGCCA	3060
Db	3001	CTCCAGACCAATTTGATCATTAACTAGTGGTGTCTGCTGCTGCACTCTCTACGCCA	3060
QY	3061	TCAACAGTGGTTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGG	3120
Db	3061	TCAACAGTGGTTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGG	3120
QY	3121	TGATTTGTTCACTTACACCAATTCCTCAAGGGTCTTATGGGAGGCAAAACCGCACCGA	3180
Db	3121	TGATTTGTTCACTTACACCAATTCCTCAAGGGTCTTATGGGAGGCAAAACCGCACCGA	3180
QY	3181	CGATTTGTCATCTCTGGGTGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3240
Db	3181	CGATTTGTCATCTCTGGGTGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3240
QY	3241	TTGATTCCTTACTACTACCGTCTCGTGGGCCCAATATCAAACTGTGGCATCAACTGCT	3300
Db	3241	TTGATTCCTTACTACTACCGTCTCGTGGGCCCAATATCAAACTGTGGCATCAACTGCT	3300
QY	3301	AGGAAAGTGGGAGTTGTAGACAGAAATATAACAGTATCGAGCAACACCGCGGA	3360
Db	3301	AGGAAAGTGGGAGTTGTAGACAGAAATATAACAGTATCGAGCAACACCGCGGA	3360
QY	3361	GCCAGAGATATTTATGTTGGGTTGTGAATTTACTAGCTTTGAGAAAGTTGTCAAAATG	3420
Db	3361	GCCAGAGATATTTATGTTGGGTTGTGAATTTACTAGCTTTGAGAAAGTTGTCAAAATG	3420

QY	3421	AGAAACACATTTGTAAATAGATGTAATAGACCATCTACGGTTTTCATGAGGTTAAGCTC	3480
Db	3421	AGAAACACATTTGTAAATAGATGTAATAGACCATCTACCGTTTTCATGAGGTTAAGCTC	3480
QY	3481	TTCTTTTGTGAAACAAAGGAATCTCATTGGTAAACCTATAGGAATTTTCCCTATAGGCA	3540
Db	3481	TTCTTTTGTGAAACAAAGGAATCTCATTGGTAAACCTATAGGAATTTTCCCTATAGGCA	3540
QY	3541	CTTTGATTGTAGGAATGGACCTATGAAATGTTGTATTTTATTATATAAAATTTATCC	3600
Db	3541	CTTTGATTGTAGGAATGGACCTATGAAATGTTGTATTTTATTATATAAAATTTATCC	3600
QY	3601	TTCTCTTACATTTTGGAGGAGTTT 3626	
Db	3601	TGTCCTTACATTTTGGAGGAGTTT 3626	
RESULT 2			
US-10-437-963-72402			
; Sequence 72402, Application US/10437963			
; Publication NO. US20040123343A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa, Thomas J.			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; APPLICANT: Wu, Wei			
; APPLICANT: Boukharov, Andrey A.			
; APPLICANT: Barbazuk, Brad			
; APPLICANT: Li, Ping			
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With			
; FILE REFERENCE: 38-21(53221)B			
; CURRENT APPLICATION NUMBER: US/10/437,963			
; NUMBER OF SEQ ID NOS: 204966			
; SEQ ID NO 72402			
; LENGTH: 3962			
; TYPE: DNA			
; ORGANISM: Oryza sativa			
; FEATURE:			
; OTHER INFORMATION: Clone ID: PAT_MBT4530_72784C.1			
US-10-437-963-72402			
Query Match 74.3%; Score 2694.4; DB 17; Length 3962;			
Best Local Similarity 88.9%; Pred. No. 0;			
Matches 2924; Conservative 0; Mismatches 361; Indels 3; Gaps 1;			
QY	55	GGCCATGACGGCGACGGGACGCCCTGAAAGTCGGGAGGACCGGGAGCGGGAGCGGT 114	
Db	368	GGCCATGACGGCGACGCCCTGAAAGTCGGGAGGACCGGGAGCGGGAGCGGT 427	
QY	115	GCCAGATCTGGCCGACGCCCTGGGCGACACGCTTGGACGGCGACGCTTTCACCGCTGCG 174	
Db	428	GCCAGATCTGGCGGACGGCGTGGGCGACGCGCGGAGGCGACGCTTTCGCGCGCTGCG 487	
QY	175	AGTCTGCGGCTTCCGGTCTGCGCGCTGCTACGAGCAGCGGCAAGAGGAGCAACC 234	
Db	488	AGTCTGCGGCTTCCGGTCTGCGCGCTGCTACGAGTACGAGCGCAAGGATGCAACC 547	
QY	235	AGGCTGCTCCAGTGCAGACCAAGTACAGCGCCACAGAGGGAGCCAGGATCCCGG 294	
Db	548	AGGCTTGGCCCCAGTGCAGACCAAGTACAAAGCGCCAAAGGGGAGCGCGGATCCCGT 607	
QY	295	GGGAGAAAGGCGACGACACTGATGCCGATGATGTTAGTGAATTCCTCACTACCTGATCTG 354	
Db	608	GGGAGAAAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 667	
QY	355	GCACTGAGGACGAGAGCAAGATTTGCTGACGATCGCAGCTGGCGCATGAACCG 414	
Db	668	GCACTGCGGACGAGAGCAAGATTTGCTGATGATGATGATGATGATGATGATGATGATG 727	
QY	415	GGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACGTGGCGAGATCGGCTCTCTCAAGT 474	

Db 728 GGGGTGGTGGAGCTCGGCGCTCCCAAGTATGACAGTGGAGATCGGCTCACCAAGT 787
Qy 475 ATGACAGTGGAGAGATCCCTTAGGGGATACGTCCTCTGAGTCAACAAGCCAGATGTGAG 534
Db 788 ATGACAGTGGGAGATCCCTCGGGGATACATCCCTTCACTCACTAATAGCCAGATCTCGG 847
Qy 535 GAGAAATCCCTGAGACTTCCCTGATCATCATGATGTCCTTACGGGAAATCATGACA 594
Db 848 GAGAAATCCCTGAGACTTCCCTGATCATCATGATGTCCTTACGGGAAATCATGACA 907
Qy 595 GAGTGTCTCGCTTCCCTATGTGAATCATTCACCAATCCGTCAAGGGAGTCTTCCGGCA 654
Db 908 AGCGTCTCAATCCCTCATGTGNAACATTCACCAATCATCAAGGGAGTCTCTGGCA 967
Qy 655 GTATGGGAATGTGGCTGGAAAGAGAGAGATTGATGGCTGGAAATGAAGCAGACAGG 714
Db 968 GCATTTGAAATGTTCCTGGAAAGAGAGATTGATGGCTGGAAATGAAGCAGACAGG 1027
Qy 715 GTGGATTCCCATGACATAATGGGACAGCATGTCTCCCTCTGAAGTTCGGGAGCTACTG 774
Db 1028 GAGCAATTCCTCATGACCAAGCAGACAGCATGTCCCTTCTGAAGTTCGGGAGTGGTG 1087
Qy 775 ACATCGATGCATCTACTGAATACAAATGGAAGACGCTTTACTGAATGATCAAACTGCC 834
Db 1088 ATATCGATGCATCCACTGATTAACAATATGGAAGATGCTTACTGAATGATGAATCGCC 1147
Qy 835 AGCCTCATCTAGAAAGTCCCATTTGCTTCTCCCAAAATTAATCCCTACAGAAATGTCA 894
Db 1148 AGCCTCTCTAGAAAAGTTCCTCTTCTCATCAGAAATAATCCCTACAGAAATGTCA 1207
Qy 895 TTGCTTCGCGTGGTGTCTAAGCATCTTCTGACACTCCGCTCTCAAAATCTCTGTC 954
Db 1208 TTGCTTCGCGTGGTGTCTAAGCAATTTCTTCTACTACCGTATTAAGATCTCTGTC 1267
Qy 955 GTAATGCATACCCACTGTGGCTTTATCTGTATATGTGAGATTGGTTTGTCTTATCCT 1014
Db 1268 GCAATCGTATCCGCTTGTATCTGTATATGTGAGATTGGTTTGTCTTGTCTCT 1327
Qy 1015 GATATCGATCAGTTCGGAAGTGGTTTCAATCAACCGGAGACCTACCTGATAGAC 1074
Db 1328 GGATATTTGGATCAGTTCGGAAGTGGTTTCAATCAACCGGAGACCTACCTGATAGG 1387
Qy 1075 TGGCTTTAAGTATGACCGAAGGTGAACCGTCTCAGTTGGCTGCTGTGACATATTG 1134
Db 1388 TGGCAITGAGTATGACAGAAAGTGAGCACTCAGTTGGCTGCTGTGACATATTG 1447
Qy 1135 TCAGTACAGTCCGCTTGAAGAGGCCACTATCGTCACTGCGCAACACTGTGCTATCCA 1194
Db 1448 TCAGTACAGTCCGCTTGAAGAGGCCACTATCGTCACTGCGCAACACTGTGCTATCCA 1507
Qy 1195 TTCTTGTGTGATATCCGTCGACAGAGTCTCTGTATGATCTGATGACGAGCTT 1254
Db 1508 TTCTTGTGTGATATCCGTCGACAGAGTCTCTGTATGATCTGATGATGGTGTG 1567
Qy 1255 CAATGCTGACTTTTGAAGCATTTGCTGAGACTTCAGAGTTTGTGAGAAATGGGTACCAT 1314
Db 1568 CAATGCTGACTTTTGAAGCATTTGCTGAGACTTCAGAGTTTGTGAGAAATGGGTCCAT 1627
Qy 1315 TTGTGAAGAGATGATGACATTGAACCCAGAGCTCCCGAGTTTACTTTTGGCAGAAATTTG 1374
Db 1628 TCGTTAAGAGATATAACATTGAGCCAGAGCTCTGATGAGTACTTCTCCAGAAATTCG 1687
Qy 1375 ATTACCTGAGAGCAAAAGTCCAGCTTCAATTTGTAAAGCCGCGGCGCATGAAGAGAG 1434
Db 1688 ATTACTTTGAAGACAAAGTCCACCTTCAATTTGTAAAGCCGCGGCGCATGAAGAGAG 1747
Qy 1435 AATATGAAGATTTAAATCAGGATTAATGTCCTAGTTTCTAAGGCAATGAAAGTCCCG 1494
Db 1748 AATATGAAGATTTCAAGTTAGGATTAATGTCCTTGTCTAAGGCAAGAAAGTCCCG 1807
Qy 1495 AGGAAGGATGATCATGCAAGATGGCACCAATGCGCAGGAAACATATCAGGAGATCATC 1554

1808 AGGAGGATGGATCATGCAAGATGGCACACCATGCGCAGGAAACAAATACTAGGACCATC 1867
1555 CTGGAATGATTCAGGTTTCCCTTGGTCAAGTGGTGGCTTGATACTAGGGTAATGAGC 1614
1868 CTGGAATGATTCAGGTTTCCCTTGGTCAAGTGGTGGCTTGATACCGGGTAATGAGC 1927
1615 TCCCGCGTTTGTGTTTATGTCTCTGTAAGAGCGTCTCTGGTTCCAGACCAACAAGAGG 1674
1928 TTCCCGCGCTGTCTACGATATCTCTGTAAGAGCGTCTCTGGTTTACAGCACCAAGAGG 1987
1675 CTGGTGCATGAATGCCCTTGTCTGTCTCAGCTGTCTTACTAATGGAATATATGT 1734
1988 CTGGTGCATGAATGTCTTGTCTGTCTCAGCTGTCTTACCAATGGAAGTACATGT 2047
1735 TGAATCTGATCTGATCTGATCACTACATCAACAACAGCAGGCTGTCCGAGAGCTATGTCT 1794
2048 TGAATCTTGTGATCTGATCACTACATCAACAACAGCAGGCTGTCCGGAAGCTATGTCT 2107
1795 TCCTAATGATCCAAACCTAGGTCGCAAGTCTGTATGTGCAAGTTCCTCAACAAGGTTG 1854
2108 TCCTTATGATCCAAACCTAGGTAAGGAGTGTCTGTATGTTCAGTTCCTCAACAAGGTTG 2167
1855 ATGGAATGATGAAGATGATCGATATGCAACAAGCAACACTGTCTTTTGTATATTAAT 1914
2168 ATGGAATGATGAAGATGATCGATATGCAACAAGCAACACTGTGTCTTTCGATATTAAT 2227
1915 TGAGGGGCTTTCAGCGCATTCAGGACCACTTTTATGTGGAACTGTGTGTCTTCAACA 1974
2228 TGAGGGGTCTTGTGATGCACTCCAAAGGACAGTTTATGTGGAACTGTGTGTCTTCAACA 2287
1975 GAACAGCTATCTATGTTTATGAGCCCCCAATTA---AGGCGAAGAGCAGGTTTCTTGG 2031
2288 GAACCTGTCTATATGTTTATGAACCCCAATTAAGCAGAAAGAGAGGGAAGTTCTTGT 2347
2032 CATCATATGTGGGGCAAGAAAGCAAGCAAGTCAAAAGAAAGAGCTCAGATAAGA 2091
2348 CATCATATGTGGGGCAGGAAGCAAGTCAAAAGAAAGAGATTGCGACAAGA 2407
2092 AAAAGTCGAACAGCATGTGCAGATCTCTCTCAGTATTCAATCTCGAAGACATAGAGG 2151
2408 AGAAGTCAAAACAGCAGTGCAGTCTGTGCGAGTTTCAATCTTGAAGATATAGAGG 2467
2152 AGGCTGTGAAGGTGCTGGGTTTGAATGAGAAATCAGTTCTCATCTCTCAATAGAGCT 2211
2468 AGGCTGTGAAGGTGCTGGATTCGATGATGAGAAATCACTTCTTATGTCTCAATAGAGCT 2527
2212 TAGAAGAGATTTGGCAGTCAAGCAGATTTGTGGCTCCACTCTGATCGAATATGCTG 2271
2528 TGGAGAAGAGATTTGGCAGTCTGAGCGTTTGTGGCTCCACTCTCATGGAATATGCTG 2587
2272 GTGTTCTCAGTCTCCACTCCAGAAATCTCTTTTGAAGAAAGCTATCCATGTCTATAAGTT 2331
2588 GTGTTCTCAATCTGCAACCCAGAAATCTCTTTTGAAGAAAGCTATCCATGTGATAAGTT 2647
2332 GTGGCTATGAGCAAGCTCTGAAATGGGAACTCAGATTTGGTTGGATCTATGGATCTGTCA 2391
2648 GTGGCTATGAGCAAGACCGAAATGGGAACTCAGATTTGGGTGGATCTACGTTTCCGCTGA 2707
2392 CAGAGATATTTCTTACTGATTCAGATGCAACGAGAGCTGCGGTTTCTAGTCTATTGCA 2451
2708 CAGAGATATTTCTTACTGATTCAGATGCAACGAGAGCTGCGGAGGCTGAGATCAATCTACTGCA 2767
2452 TGCCCAAGCGCCAGCTTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATGCTCTGAACC 2511
2768 TGCCCAAGCGCCAGCTTTTCAAGGGGTCTGTCTCTATCAATCTTTTCAGATGCTCTTAACC 2827
2512 AAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGCGCGCATTTGCCCTTAT 2571
2828 AAGTGTTCGGTGGGCACTTGGTTCTGTGAAATTTCTTTTTCAGTCGCAATTTGCCCATAT 2887
2572 GGTATGCTCAGGAGGCGCTCAAGTTCTCTGAGAGATTCGCTTACATCAACACCACA 2631
2888 GGTACCGCTATGGAGGAGCGCTTAAAGTTCTTGGAGAGATTTGCCCTATCATCAACACCACA 2947

```
QY 2632 TTTACCACTAACCTCTCTCCGCTCTAGTCTATGTATATGCTGCTATCTGTCTGC 2691
Db 2948 TTTATCCATTGACATCGATCCGCTCTCTCATATACGTGTTTGGCTGTCTGTCTGC 3007
QY 2692 TCACCTGGAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCATTGCGC 2751
Db 3008 TCACTGGGAAGTTTCATCATGCCAGAGATTAGCAACTTGGCTAGTATTTGGTTCACTCTC 3067
QY 2752 TCTTCTCTTCAATTTTGGCCATCTGTTATCCTTGAGATGAGGTGGAGTGTGGCATGG 2811
Db 3068 TCTTCTATTTCAATTTTGGCCATCTGTTATCCTTTGAGATGAGGTGGAGTGTGGCATCG 3127
QY 2812 ACGAGTGGTGGAGGAATGAACAGTTTCTGGGTCTATTGGAGGTATCTCTGCACATCTGTTTG 2871
Db 3128 ATGAGTGGTGGAGGAATGAACAGTTTCTGGGTATTGGAGGTATCTCTGCACATCTTCTTG 3187
QY 2872 CCGTCTTTTCAGGTTCTTCTGAAGGTTCTTGGCGGTATCGACCACTTCACTGTCACT 2931
Db 3188 CCGTCTTTCAGGTTCTTCTGAAGGTTCTTGGCTGTATCGACCACTTCACTGTCACT 3247
QY 2932 CAAAGGCTTAATGACGAAGAGGCACTTTGCTGAGCTTACATGTTCAAGTGGACGACGC 2991
Db 3248 CAAAGGCTTCTGATGAAGATGGCACTTTGCTGAGCTTACATGTTCAAGTGGACGACGC 3307
QY 2992 TTCTCATCTCCCGACGACCATTTTGTATCATTAACATGTTGTTGTTGTTGTTGTTGTTG 3051
Db 3308 TTCTCATCTCCCGACGACCATTTTGTATCATTAACATGTTGTTGTTGTTGTTGTTGTTG 3367
QY 3052 CTTAGCCCATCAACAGTGGTTACCAATCATGCGGCGCTCTTTGGGAAGCTCTTCTTTTG 3111
Db 3368 CATACCGCATCAACAGCGCTACCAATCATGCGGCGCTCTTTGGGAAGCTCTTCTTTTG 3427
QY 3112 CTTTCTGGGTGATTTGTTACATTATACCACTTCTCAAGGTTCTTATGGCAGGCAAAACC 3171
Db 3428 CTTTCTGGGTGATTTGTTACATTATACCACTTCTCAAGGTTCTTATGGCAGGCAAAACC 3487
QY 3172 GCACACCGACGATTCTCATGCTGGCTGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 3231
Db 3488 GCACCTCGACCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3547
QY 3232 GGGTTCTGTTGATCCATTCACTACCGCTCTGCTGCGCCCAATATCCAAACCTGTGGCA 3291
Db 3548 GGGTTCTGATCGATCCATTCACTACCGCTCTGCTGCGCCCAATATCCAAACCTGTGGCA 3607
QY 3292 TCAACTGCTAGGAAGTGGAGTTTGTAGACACAGAAATATACAGT 3339
Db 3608 TCAACTGCTAGGAAGTGGAGTTTGTAGACACAGAAATATCTGCCAGT 3655
```

RESULT 3

```
US-10-627-132-21
; Sequence 21, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE OF INVENTION: Theof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
```

; LENGTH: 3799

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 3757..3775, 3777, 3782

; OTHER INFORMATION: n = A,T,C or G

US-10-627-132-21

Query Match

71.6%; Score 2595; DB 13; Length 3799;

Best Local Similarity 87.1%; Pred. No. 0;

Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;

QY 36 CCGGTTGGATCGCTCTGCGCGCCATGACGCGGAGCGGAGCCCTGAAGTCCGGGAGG 95

Db 214 CCGTTCTCGGCGAGCTCGCTTCCATGAGGAGCGAGCGGCGTGAAGTCGGGAGG 273

QY 96 CACGGGGCGGGAGACGTGTGCCAGATCTCGCGGACGGCTCGGCACACACGTTGACGCG 155

Db 274 CCGGGGGAGGCGAGGTGTGCCAGATCTCGGCGGATGGCGTGGGCATACGCGGGAGGA 333

QY 156 GACGTCTTACCGCTCGGACGTCTGCCGCTTCCCGGCTTGC CGCCCTGTCTACGAGCAC 215

Db 334 GACGTCTTACCGCTCGGACGTCTGCCGCTTCCCGGCTTGC CGCCCTGTCTACGAGTAC 393

QY 216 GAGCGAAGGAGGCGACCGAGCCTTCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 275

Db 394 GAGCGAAGGAGGCGACCGAGCCTTCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 453

QY 276 GGGAGCCAGCGATCCGCGGGGAGGAGGAGCGACGACACTGTATGCCGATGATGATGATGAC 335

Db 454 GGGAGTCCAGCGATCCGCGGGGAGGAGGAGCGATGATGATGCCGATGATGATGATGATGAC 513

QY 336 TTCAACTACCTGTCAATCTGGCACTGAGGACCAAGAGAGAGAGATTGCTGACAGGATGCGC 395

Db 514 TTCAACTACCTGTCAATCTGGCAATGACGACCAAGAGAGAGATTGCTGACAGGATGCGC 573

QY 396 AGCTGGCGCATGAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGC 455

Db 574 AGCTGGCGCATGAATGCTTGGGGGAGCGGGGATGTTGGCCGCGCCCAAGTATGACAGTGGT 633

QY 456 GAGATCGGCTCTCCAACTATGACAGTGGAGAGATCCCTAGGGGATAGTCCCTTCATGTC 515

Db 634 GAGATCGGCTTACCAAGTACGACAGTGTGAGATCCCTTCGGGGATATATCCCTTCATGTC 693

QY 516 ACCAAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCCCTGATCATCATGATGATGTC 575

Db 694 ACTAAGCCAGATTTTCGGGAGAAATCCCTGGTGTCTCCCTGACCATCATGATGATGTC 753

QY 576 CTTAGCGGGAACATCAGCAGAGTGTCTCGTTTCCCTATGATCATTCACCAATCCG 635

Db 754 CTTAGTGGGAACATTTGGCAGGCGCGCCCATTTCCCTATATGAATCATTCATCAATCCG 813

QY 636 TCAGGGAGTTCTCCGCGAGTATGGGAATGTTGCTCGAAAGAGAGAGATTGATGGCTGG 695

Db 814 TCGAGGGAATTTCTCTGGTAGGTTGGGAATGTTGCTCGAAAGAGAGAGATTGATGGCTGG 873

QY 696 AAAATGAAGCAGGACCAAGGTTGCCATTCCTATGATTAATGGGACCAAGCATTTGCTCCCTCT 755

Db 874 AAAATGAAGCAGGACCAAGGGAACAATTTCCCATGCAATGGCAGATGCAAGCATTTGCTCCCTCT 933

QY 756 GAAAGTGGGCGAGCTACTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 815

Db 934 GAGGCGCGGGTGTGGTGTATTTGATGATCAACTGATTAACCAATGATGATGATGATGATGAT 993

QY 816 CTGAATGATGAACCTCGCCAGCTCTATCTAGAAAGTCCCATTTGCTTCCTCCCAATA 875

Db 994 TTAACGATGAACCTCGCCAGCTCTATCTAGAAAGTTCACCTTCCTTCCTCCAGGATA 1053

QY 876 AATCCCTACAGATGGTCAATTTCTGCGGTTGGTGTGTTCTTAAGCATCTTCTGCTACCTAC 935

Db 1054 AATCCATACAGATGGTCAATTTCTGCTAGATTTGATTTGTTCTTAAGCATCTTCTGCTACCTAC 1113

Db 3271 GGCAAGCTCTTCTTCGCTCTCTGGGTGATGCTTACCTGTATACCCGTTCTCCTCAAGGCTCTC 3330
QY 3156 ATGGCAGGCAAAACCGCACACGAGGATTGTCATCGTCTGGGCTGTCTCTCTCGCTTCT 3215
Db 3331 ATGGGTGCGCAAGAACCGCACCCCGACCATCGTGGTGTGCTGGCGGATCTCTGGCGTCG 3390
QY 3216 ATCTTCTCTCTGCTGCTGGGTGCTGTGATCATCTTCACTTACCCGTTCTCGTGGCCCAAT 3275
Db 3391 ATCTTCTCTGCTGTGGTTCGATCGATCGTTCCACACCGCGTCACTGGCCCGGAT 3450
QY 3276 ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGA-GTTGTAGACAGAAA 3329
Db 3451 ACTCGAAGCTGTGGCATCAACTGCTAGGAGGTGGAAGTTGTAGAAACAGAGA 3505

RESULT 4
US-10-267-459-5
; Sequence 5, Application US/10267459
; Publication No. US20030150014M1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864RC
; CURRENT APPLICATION NUMBER: US/10/267,459
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3799
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)...(3477)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(3799)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-459-5

Query Match 71.6%; Score 2595; DB 15; Length 3799;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;

QY 36 CGGTTTGATCGCTCTGCGGCCATGACCGCGACGCGGACGCGCTGAAGTCCGGGAGG 95
Db 214 CGGTTCTGGCGAGCTCGCTCGCCATGAGGCGGACGCGGCGTGAAGTCCGGGAGG 273
QY 96 CACGGGCGGGAGCGTGTGGCAGATCTGCGCCGACGCGCTGGGACACACGTTGGAACGC 155
Db 274 CGCGGGGAGGGCAGGTGTGCAGATCTGCGCGGATGCGTGGGCATCTACCGCGGAGGA 333
QY 156 GACGCTTTACCGCTGCGAGCTCTGCGCTTCCCGCTTCCCGCTCTGCGCCCTGCTACGAGCAC 215
Db 334 GACGCTTTACCGCTGCGAGCGTGTGCGGTTTCCCGGTGTGCGCGCCCTGCTACGAGTAC 393
QY 216 GAGCGCAAGGAGGCGACCCAGGCTGCTTCCAGTSCAAGACCAAGTACAAGCGGCCACAGA 275
Db 394 GAGCGCAAGGAGGCGCACAAAGCGTGCCTGAGTGCACCAAGTACAAGCGGCCACAG 453
QY 276 GGGAGCCAGCATCCGCGGGGAGGAGGACGACGACCTGATCGCGATGATGTAGTAC 335
Db 454 GGGAGTCCAGCGATCCGAGGGGAGGAGGACGATCTGATCGCGATGATGTAGTAC 513
QY 336 TTCAACTACCTGCTGCTGCGACTTGAGGACCGAGAGGAGATTGCTGACAGGATGCGC 395

Db 514 TTCAACTACCTGCTGCTGCGCAATGACGACCGAGAGAGATTGCTGACAGGATGCGC 573
QY 396 AGCTGGCGCATGAACACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGC 455
Db 574 AGCTGGCGCATGAATGCTGGGGGCGAGCGGGGATGTTGGCGCGCCCAAGTATGACAGTGGT 633
QY 456 GAGATCGGCTCTCCAGTATGACAGTGGAGAGATCCCTAGGGGATAGTCCCTTCACTCAGTC 515
Db 634 GAGATCGGCTCTTACCAAGTACGACAGTGGTGGAGATCCCTCGGGGATATATCCCGTCACTC 693
QY 516 ACCAACGCGCAGATGTTCAGGAGAAATCCCTGGAGCTTCGCTGTGATCATCATGATGTCC 575
Db 694 ACTAACGCGCAGATTTCCGGGAGAAATCCCTGGTGTCTCCCTGACCATCATATGATGTCT 753
QY 576 CCTACGGGGAACATCAGCAGACGTGCTCCGTTTCCCTATGTGAATCATTCACCAATCCG 635
Db 754 CCTACTGGGAACATTTGGCAGGCGGCGCCATTTCCCTATATGAATCATTCATCAATCCG 813
QY 636 TCAAGGGAGTTTCTCGGCGAGTATTGGGAATGTTGGCTGGAAAGAGAGAGATTGATGGCTGG 695
Db 814 TCGAGGGAAATTTCTCTGCTAGCGTTGGGAATGTTGCCCTGGAAAGAGAGGTTGATGGCTGG 873
QY 696 AAAATGAAGCAGGACCAAGGCTGCCGATTCCTCATGACTAATGGACAAAGCATTCCTCCCTCT 755
Db 874 AAAATGAAGCAGGACCAAGGGAACATTTCCCATGACGAATGGACAAAGCATTCCTCCCTCT 933
QY 756 GAAAGTGGGCGAGCTACTGACATCGATCATCTGAATCAACATGAGAAAGAGAGAGCTTTA 815
Db 934 GAGGCGCGGGGTGTTGGTGATATTGATGATCACTCACTGATTACAACTGAGAAAGTGCCTTA 993
QY 816 CTGAATGATGAACCTCGCCAGGCTCTATCTAGAAAGTCCCATGCTTCTCCAAAATA 875
Db 994 TTAACCGATGAACCTCGCCAGGCTCTATCTAGAAAGTTCACATTCCTTCTCCAGGATA 1053
QY 876 AATCCCTACAGAATGGTCAATGTTCTGCGGTGTTGTTCTTAAGCATCTTCTCGCACTAC 935
Db 1054 AATCCATACAGAGTGGTCAATGTTGCTAGCATGATTGTTCTAAGCACTTCTGACACTAC 1113
QY 936 CGTCTCAAAATCCTGTGCGTAATGCATACCACCTGTGGCTTTTATCTGTTATATGTAG 995
Db 1114 CGGATCAAAATCCTGTGCGTAATGCATACCACCTGTGGCTTCTATCTGTTATATGTAG 1173
QY 996 ATTTGGTTGCTTTTATCTGATGATCTGATCAGTTCGGAAGTGGTTTCCATCAGACGG 1055
Db 1174 ATCTGGTTGCTCTTTCTGATATTTGGATCAGTTTCCAAAGTGGTTTCCAAATCAACCGC 1233
QY 1056 GAGACCTACCTTGATAGACTCGCTTTAAGGTATGACCGGAGAGGTGAACCGTCTCAGTTG 1115
Db 1234 GAGACTTACCTTGATAGACTCGCATTAAGGTATGACCGGAGAGGTGAGCCATCTCAGTTG 1293
QY 1116 GCTGCTGTGACATATTTGTGTCAGTACAGTCAACCCCTTGAAGGAGGCCACTATGCTACT 1175
Db 1294 GCTGCTGTGACATATTTGTGTCAGTACAGTCAACCCCTTGAAGGAGGCCCTCTCTGCTACT 1353
QY 1176 GCCAAGCTGTGCTATCCATTTCTGCTGTTGATTTATCCGTTGGACAGGTTCTTGTCTAT 1235
Db 1354 GCCAATACCGTGTATCCATTTCTGCTGTGAGTATCTGCTGGAATAGGTTCTTGTCTAT 1413
QY 1236 GTATCTGATGACGAGCTTCAATGCTGACTTTTTCAGCATTTGGCTGAGACTTTCAGAGTTT 1295
Db 1414 GTATCTGATGATGAGCTGTCTATGCTGACATTTGATGCACTAGCTGAGACTTTCAGAGTTT 1473
QY 1296 GCTAGGAAATGGGTACCATTTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTT 1355
Db 1474 GCTAGAAAATGGGTGCCATTTTGTGAAGTATCAACATTTGAACCTAGAGCTCCCTGAATGG 1533
QY 1356 TACTTTTCCAGAAAATTTGATCTCTGAAAGACAAAGTCCAGCCCTCATTTGTTTAAAGAC 1415
Db 1534 TACTTCTCCAGAAAATTTGATTTACTTGAAGGACAAAGTGCACCCCTCATTTGTTTAAAGAC 1593
QY 1416 CGCCGGGCGCATGAAGAGAAATATGAAGAATTTTAAATCAGATAAATGCCTAGTTTCT 1475

Db 1594 CGCGGGCCATGAAGAGAGATATGAAGAATTCAAATTTAGGGTAAATGGCCCTTGTGCT 1653
Qy 1476 AAGGCATTGAAGTCCCAGGAGAGATGATCATCAAGATGCGACACCATGGCCAGGA 1535
Db 1654 AAGGCACAAAAAGTCCCTGAGGAAGATGATCATCAAGATGCGACACCATGGCCAGGA 1713
Qy 1536 AACAAATACCAAGGATCATCTCGAATGATTCAGGTTTTCCTGGTCACAGTGTGGCCCTT 1595
Db 1714 AACAAATACCAAGGACCATCTCGAATGATTCAGGTTTTCCTGGTCACAGTGTGGCTT 1773
Qy 1596 GATACAGAGGTAATGAGCTCCCCGTTTATGTTATGTCTCGTGAAGAGCGTCTGGG 1655
Db 1774 GATACAGGTAATGAGCTACCCGTTTGTCTATGTTCTCGTGAAGAGCGTCTGGG 1833
Qy 1656 TTCCAGCACCAAGAGGCTGTGCGATGAATGCTTGTCTCGTCTCAGCTGTCCCT 1715
Db 1834 TTCCAGCATCACAAAGAGCTGTGCGATGAATGCTTGTCTCGTCTCAGCTGTCCCT 1893
Qy 1716 ACTAATGGCAATACATGTTGAATCTTGTATGTGATCACTACATCAACACAGCAAGGCT 1775
Db 1894 ACCAATGGCAATACATGTTGAATCTTGTATGTGATCACTACATCAACACAGTAAGGCT 1953
Qy 1776 GTCGAGAGAGCTATGCTTCTTAATGATCCAAACCTAGGTCGCAAGTCTCTTATGTG 1835
Db 1954 CTGAGGAAAGCTATGCTTCTTATGGATCTTAACCTAGGAAGGAGTGTCTGCTATGTT 2013
Qy 1836 CAGTCCCAAGAGGTTGATGGATGTAGGAATGATGAAATGATGATCAACACAGCAAGCACT 1895
Db 2014 CAGTTCCTCCAGAGGTTGATGGATGTAGGAATGATGAAATGATGATCAACACAGCAAGCACT 2073
Qy 1896 GTCTTTTTCATATTAACCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 1955
Db 2074 GTCTTTTTCATATTAACCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 2133
Qy 1956 ACTGGTTGTGTTTCAACAGAAACAGCTATCTATGTTATGAGCCCAATTAAGGCGAAG 2015
Db 2134 ACTGGCTGTGTTTCAACAGAAACAGCTATCTATGTTATGAGCCCAATTAAGGCGAAG 2193
Qy 2016 AAGCAGGTTCTTGGCATCACTATGTTGGGGCAAGAGAGCAAGCAAGTCAAGAGAA 2075
Db 2194 AAGGGTGTGTTTCAACAGAAACAGCTATCTATGTTATGAGCCCAATTAAGGCGAAG 2252
Qy 2076 AGAGCTCAGATAAGAAAAGTCCAAACAGCATGTGACAGTGTCTGTTCCAGTATTCAAT 2135
Db 2253 --GGGCTCAGACAGAAAAGTCCAAACAGCATGTGACAGTGTCTGTTCCAGTATTCAAT 2310
Qy 2136 CTCGAAGACATAGAGAGGCTGTGAAGTGTGAGTGTGAGTGTGATGATGAGAAATCAGTTCTC 2195
Db 2311 CTTGAAGATATAGAGGAGGAGTGAAGGCTGTGATTTGATGATGAGAAATCAGTTCTT 2370
Qy 2196 ATGTCTCAATGAGCTTAGAGAGAGATTTGGCCAGTCAGCAGCATTTGTTGCTCCACT 2255
Db 2371 ATGTCTCAATGAGCTTAGAGAGAGATTTGGCCAAATCTGACGCTTTGTTGCGTCCACT 2430
Qy 2256 CTGATGGAATATGCTGAGTGTCTCAGTCCCTCAGATCCAGAAATCTCTTTTGAAGAAGCT 2315
Db 2431 CTGATGGAATATGCTGAGTGTCTCAGTCCCTCAGATCCAGAAATCTCTTTGAAAGAAGCT 2490
Qy 2316 ATCCATGTCAATAGTGTGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTGG 2375
Db 2491 ATCCATGTCAATAGTGTGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTGG 2550
Qy 2376 ATCTATGATCTCTCAGAGATATTTCTTACTGATTCAAGATGACGCAAGAGGCTGG 2435
Db 2551 ATCTATGATCTCTGAGCGAAGATATTTCTCAGTGGGTTCAAGATGACGCAAGAGGCTGG 2610
Qy 2436 CGTTCAGTCTATTTGATGCCCAAGCGCCAGCTTTTCAAGGATCTGCCCCCAATCAATCTT 2495
Db 2611 CGGTGATCTACTGATGCTTAAAGCGCCGCGCTTCAAGGATCTGCGTCCCAATCAATCTC 2670
Qy 2496 TCAGATGCTGACCAAGTGTGCTGCGGTGGGCTCTGGTCTGTTGTAATCTTTTCAGC 2555
Db 2671 TCAGACCGTCTGAACCAAGTGTGCTGCGGTGGGCTCTGGTCTGTTGTAATCTTTTCAGC 2730

Qy 2556 CGGCATTGCCCTTATGGTATGGCTACGAGGCGCCTCAAGTTCCTGGAGAGATTGCT 2615
Db 2731 CGGCATTGCCCTTATGGTATGGCTACGAGGCGCCTCAAGTTCCTGGAGAGATTGCT 2790
Qy 2616 TACATCAACACCAACCAATTTACCACTAACCTCTCCCGCTTCTAGTCTATTTGATATTG 2675
Db 2791 TACATCAACACCAACCAATTTACCACTAACCTCTCCCGCTTCTAGTCTATTTACCTCTG 2850
Qy 2676 CTTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2735
Db 2851 CTTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2910
Qy 2736 ATCTGTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2795
Db 2911 ATCTGTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2970
Qy 2796 AGTGGTGTGGCAATGACAGTGTGGAGGAATGAACAGTCTTGGGTCTATTTGAGGTATC 2855
Db 2971 AGCGGCTGGGATGACAGGAGTGTGGAGGAACGAGCAGTCTTGGGTCTATCGGAGGCATC 3030
Qy 2856 TCTGACATCTGTTTGGCGCTTTCAGGCTCTTCAAGGTCTTCAAGGTCTTCCCGTATCGACACC 2915
Db 3031 TCGGCCACCTCTTTCAGGCTCTTCCAGGCTCTTCAAGGTCTTCCCGTATCGACACC 3090
Qy 2916 AACTTCACCTGTCACTTCAAGGCTAATGACGAAGAGGCACTTTCGCTGAGCTCTACATG 2975
Db 3091 AACTTCACCTGTCACTTCAAGGCTCTGATGAGAGCGGCACTTTCGCGAGCTGTACATG 3150
Qy 2976 TTCAAGTGAAGCAGCTTCTATCTTCCGACGACCAATTTTGTATCAATTAACATGTTGGT 3035
Db 3151 TTCAAGTGAAGCAGCTTCTATCTTCCGACGACCAATTTTGTATCAATTAACATGTTGGT 3210
Qy 3036 GTCTGTTGCTGGCAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 3095
Db 3211 GTTGTGCGGCACTTCTACGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 3270
Qy 3096 GGAAGCTCTTCTTTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 3155
Db 3271 GGAAGCTCTTCTTTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 3330
Qy 3156 ATGGGAGGAAACCGGACCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3215
Db 3331 ATGGGAGGAAACCGGACCGGACCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3390
Qy 3216 ATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3275
Db 3391 ATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3450
Qy 3276 ATCCAAACCTGTGGCAGTCACTGCTAGGAAAGTGGGA-GTTTGTAGAGACAGAAA 3329
Db 3451 ACTCGAAGCTGTGGCAGTCACTGCTAGGAGGTGGAAAGTTCGTAGAAACAGAGA 3505

RESULT 5

US-10-209-059-21
; Sequence 21, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 3799
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 3757..3775, 3777, 3782
; OTHER INFORMATION: n = A,T,C or G
US-10-209-059-21

Query Match
Best Local Similarity 87.1%; Pred. No. 0;
Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;

QY 36 CGGGTTGGATCGCTCTGCCGCGCCATGAGCGGGACGCGGAGCCCTGAAGTCCGGGAGG 95
DB 214 CGGTTCTGGCGGAGCTCGCTCCATCGGAGGCGACGCGGCTGAAGTCGGGGAGG 273
QY 96 CACGGGGCGGGAGCCTGTGCCAGATCTCGCGGACGGCTGGGCACCACTGTGGACGGC 155
DB 274 CGCGGGGAGGCGAGGTGTGCCAGATCTCGGGGATGGCGTGGGCACTACGGCGGAGGA 333
QY 156 GACGCTTTCACCGCTGCGACCTGTGCGGCTTCCCGGTCTGCGGCCCTCTGCTACGAGCAC 215
DB 334 GACGCTTTCACCGCTGCGACCTGTGCGGCTTCCCGGTCTGCGGCCCTCTGCTACGAGTAC 393
QY 216 GAGCCCAAGGAGGCGACCCAGCGCTCTCCAGATCTCGCGGACGGCTGGGCACCACTGTGGACGGC 275
DB 394 GAGCCCAAGGAGGCGACCAAGCGCTGCCCGCCAGTGCAGAAACAAAGTACAAAGCGCCACAAAG 453
QY 276 GGGAGCCAGCGATCCGCGGGGAGGAGCGGACGACACTGATGCCGATGATGCTAGTGCAC 335
DB 454 GGGAGTCCAGCGATCCGAGGCGGAGGAGAGGATCTGATGCCGATGATGCTAGCGGAC 513
QY 336 TTCAACTACCTTGCATCTGGCACTGAGGACCAAGAGCAGAAATGCTCAGAGGATGCGC 395
DB 514 TTCAACTACCTTGCATCTGGCAATGACGACCAAGAGCAGAAATGCTCAGAGGATGCGC 573
QY 396 AGCTGGCGCATGAACACCGGGGCGAGTGCATGTTGGCCACCCAGATGACAGTGCAGAGTGGC 455
DB 574 AGCTGGCGCATGAATGCTGGGGGCGAGCGGGGATGTTGGCGGCCCAAGATGACAGTGGT 633
QY 456 GAGATCGGCTCTCAAGTATCACAGTGGAGATCCCTAGGGGATAGCTCCCTTCAGTC 515
DB 634 GAGATCGGCTTACCAAGTACACAGTGTGAGATCCCTCGGGATACATCCCTCGTACGTC 693
QY 516 ACCAAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCACATGATGTC 575
DB 694 ACTAAGCCAGATTCGGGAGAAATCCCTGGTCTTCCCTTGACCATCATGATGTC 753
QY 576 CCTACGGGACATCAGCAGAGTCTCGGTTCCCTATGTAATCATTCACCAATCGG 635
DB 754 CCTACTGGAAATGGCAGGCGCGCCCATTCCTCTATGAAATCATTCATCAAAATCGG 813
QY 636 TCAAGGGAGTCTCCCGGAGTATGGGAATGTGCTGGAAAGAGAGATTTCATGGCTGG 695
DB 814 TCGAGGGAATCTCTGTTAGCTTGGGATGTGCTGGAAAGAGAGGTTGATGGCTGG 873
QY 696 AAAATGAAGCAGGACAAGGGTGCATCCCATGACTAATGGGACAAAGATTCCTCCCTCT 755
DB 874 AAAATGAAGCAGGACAAGGGAAATTCCTCATGCGAATGGCACAAGCATTCCTCCCTCT 933
QY 756 GAAGTCCGGGAGCTACTGCACATCATGATCTACTGATACAACTGAGAGAGCGCTTTA 815
DB 934 GAGGCGCGGGGTGTGGTGATATGATGATCACTCAACTGATTAACAATGGAAGATGCCTTA 993
QY 816 CTGAATGATGAACCTCGCCAGCTCTATCTAGAAAGTCCCATGCTTCCTCCAAATA 875
DB 994 TTAACAGATGAACCTCGCCAGCTCTATCTAGAAAGTTCACATTCCTTCCTCCAGGATA 1053
QY 876 AATCCCTACAGAAATGGTCAATGTTCTGCGGTTGGTGTCTTAAGCATCTTCCTGCACTAC 935
DB 2134 ACTGGCTGTGTTTCAACAGAAACAGCTCTATGTTGTTATGAGCCCCCANNATTAGCAAAAG 2193
```



```
QY 2016 AAGCAGGTTCTTGGCATCACTATCTGGGGCAAGAGGCAAGCAAGTCAAGAAA 2075
Db 2194 AAGGGTGGTTCTTGTATCACTATCTATGTGGGAGGAAGGAGCAAAATCAAGAA- 2252
QY 2076 AGGAGCTCAGATAAGAAAAGTCCGAACAAGCATGTGGACAGTTCTGTCTTCAGTATTCAAT 2135
Db 2253 --GGGCTCAGACNAGAAAAGTACAGNAGCATGTGGACAGTTCTGTGCCAGTATTCAAT 2310
QY 2136 CTCGAAGACATAGAGAGGGTGTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTC 2195
Db 2311 CTTGAAGATATAGAGAGGGAGTTGAAGGCGCTGGATTTGATGATGAGAAATCACTCTT 2370
QY 2196 ATGCTCAATGAGCTTAGAGAGAGATTGGCCAGTACAGCAGATTGTGCTCCACT 2255
Db 2371 ATGCTCAATGAGCTTAGAGAGAGATTGGCCAAATCTCAGCTTTGTGGTCCACT 2430
QY 2256 CTGATGAATPATGGTGTGTTCTCCTCAGTCTCCACTCCAGAAATCTCTTTTGAAGAAGCT 2315
Db 2431 CTGATGAATPATGGTGTGTTCTCCTCAGTCTCGACTCCAGAAATCTCTTCTGAAAAGAGCT 2490
QY 2316 ATCCATGCTAATAGTTGTGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTGG 2375
Db 2491 ATCCATGCTAATAGTTGTGCTACGAGGACAAAGATTGAATGGGAACTGAGATTGGTGG 2550
QY 2376 ATCTATGGATCTGTACAGAGATATTCTTACTGGATTCAAGATGCAAGCAAGAGGCTGG 2435
Db 2551 ATCTATGGTCTGTGACGAGAGATATTCTCACTGGGTTCAAGATGCAAGCAAGAGGCTGG 2610
QY 2436 CGTTCAGTCTATTGTCATGCCCAAGCGCCAGCTTTTCAAGGATCTGCCCCCAATCAATCTT 2495
Db 2611 CGTTCGATCTACTGCTATGCTTAAGCGCGCGCTTCAAGGATCGGCTCCCATCAATCTC 2670
QY 2496 TCAGATGCTTGAAACAAAGTCTCGGTGGGCTCTGGTCTGTTGAAATCTTTTCAGC 2555
Db 2671 TCAGACGCTGTGAACCAAGTGTCTCGGTGGGCTCTCGGTTCAAGTGAATCTTTTCAGC 2730
QY 2556 CGCATTTGCCCTTATGCTATGCTACGAGGGCGCTCAAGTTCTCTGGAGAGATTCCGCT 2615
Db 2731 CGCATTTGCCCTTATGCTACGAGGTCAGAGGACGCTCAAGTTCTCTGGAGAGATTCCGCT 2790
QY 2616 TACATCAACACCACTATTACCACTAACCTCTCTCCCGTTCTPAGTCTATTGATATG 2675
Db 2791 TACATCAACACCACTATTACCCGCTCACGCTCCCTCCGCTCTCTCAATTTACTGTATCCTG 2850
QY 2676 CTTGCTATCTGTGCTCACTGNAAGTTCATATGCCAGAGATTACGAACTTGGCCAGT 2735
Db 2851 CTTGCCATCTGCTGCTCACGGGAAGTTTCATATCCAGAGATCAAGAACTTCGCTAGT 2910
QY 2736 ATCTGGTTTCAATGGCTCTTCTTCAATTTTCGCCACTGCTATCTTGAAGATGAGTGG 2795
Db 2911 ATCTGGTTTCAATCTCTTCAATCTCGATCTTCCGCAAGGATCTCTGGAGATGAGTGG 2970
QY 2796 AGTGGTGTGGCATGTGAGAGATGAACAGTTCTGGGTCTATTGAGGTATC 2855
Db 2971 AGCGCGTGGGCATCGACGAGTGTGGAGGAACGAGCAGTTCTGGGTCTATCGGAGGATC 3030
QY 2856 TCTGCATCTGTTTGGCTTTTCAGGTTCTTCAAGTGTCTGAGGTGCTTGGCGGTATCGACCC 2915
Db 3031 TCGGCCACCTCTTTCGGCGCTTCTCCAGGGCTCTCTCAAGGTGCTTGGCGGATCGACCC 3090
QY 2916 AACTTCACTGTCACTCAAGGCTTAATGAACGAAGGCGACTTTGCTGAGCTCTACATG 2975
Db 3091 AACTTCACTGTCACTCAAGGCTCTCGATGAAGACGCGACTTTCGGGAGCTGTACATG 3150
QY 2976 TTAAGTGGACGAGCTTCTCATCCCTCCGACGACCAATTTTGATCATTAATGTTGGT 3035
Db 3151 TTAAGTGGACGACACTTCTGATCCCGCCCAACCACTCTGTATCATCAACCTGGTCGGC 3210
QY 3036 GTCGTCTGCGACCTCTACGCCATCAACAGTGGTTACCAATCATGGGGCGCTCTTT 3095
Db 3211 GTTGTTCGGGCACTCTCTACGCCATCAACAGCGGGTACCAGTCTGGGGTCCGCTCTTC 3270

RESULT 6
US-10-160-719-29
; Sequence 29, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321) ... (3551)
US-10-160-719-29

Query Match 70.9%; Score 2572.6; DB 15; Length 3746;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

QY 4 CGAGGAACCCCGCTCCAGCTCTCTGTCGGTGGGTTGGATCGCTCTGCCGCCATGG 63
Db 265 CCAGGCCCGAGGCTCCAGGCCAGTCCCTCGACGTTTCTCGGGAGCTCGCTTGCATGG 324
QY 64 ACGGCGACGCGAGCCCTCGAAGTCCGGGAGGCAACGGGGCGGGGACGTGTGCCAGATCT 123
Db 325 AGGCGCACGCGGACGCGGTGAAGTTCGGGGAGGCGCGGTGGCGGACAGTGTGCCAGATCT 384
QY 124 GCGCGACGCGCTGGGACCAACGTTGGACGCGGACGCTTTCACCGCTCGGACGCTTGGCC 183
Db 385 GCGCGACGCGGTGGGACCAACGCGGAGGGGAGACGTCTTCGCCGCTTCGACGCTTGGCG 444
QY 184 GCTTTCGCGTCTGCCGCCCTTGTCTACGACACGAGCGCAAGAGGGGACCCAGGCGCTGCC 243
Db 445 GGTTTCGCGTGTGCCGCCCTTGTCTACGAGTACGAGCGCAAGGACGCGCACGCGGCGTGGC 504
QY 244 TCAGTTCGAAGCAACAGTACAGCGCCACAGAGAGGCCAGCGATCCGGGGAGGAG 303
Db 505 CCCAGTTCGAAGCAACAGTACAGCGCCACAGGGGAGGCCCGGCGATCCGTTGGGAGGAG 564
QY 304 GCGACGACACTGATCGCGGATGATGGTAGTGACTTCAACTACCTGCTATCTGGCACTGAGG 363
Db 565 GAGACGACACTGATGCCGA-----TAGGACCTTCAATTACCTTGCATCTGGCAATGAGG 618
```


QY 364 ACCAGAAGCAGAGATTTGCTGACAGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTG 423
Db 619 ACCAAGACGAAGATTTGCCGACAGATGCGCAGCTGGCGCATGAACCTGGGGGCGAGG 678
QY 424 GCAATGTTGGCCACCCCAAGTAGTACAGTGGGAGATCGGCCCTCTCCAAGTAGACAGTG 483
Db 679 GGGATGTTGGTCGCCCCCAAGTAGTACAGTGGCGAGATCGGGCTTACCAAGTAGACAGTG 738
QY 484 GAGAGATCCCTAGGGGATAGTCCCTTCTAGTCAACCAAGCCAGATGTCAGGAGAAATCC 543
Db 739 GCGAGATTCCTCGGGGATACATCCCATCAGTCAGTCACTAAACAGCCAGATCTCAGGAGAAATCC 798
QY 544 CTGGAGCTTCGCTGATCATCATCATGATGTCCTCCCTACGGGGAACATCAGCAGCGTCTC 603
Db 799 CTGGTGTCTCCCTGACCATCATATGATGTCCTCCCACTGGGAACATGGCAAGCGTGTCTC 858
QY 604 CTTTTCCCTATGTGAATCATTCACCAAAATCCGTCAAGGGAGTTCCTCCGCGCATTTGGGA 663
Db 859 CATTTCCCTATGTGAACCATTCGCCAAATCCGTCAAGGGAGTTCCTGGTAGCATTTGGGA 918
QY 664 ATGTTGCCCTGGAAGAGAGAGATGTAGTGGCTGGAAATGAAGCAGGACGAAGGTGGATTC 723
Db 919 ATGTTGCCCTGGAAGAGAGAGGTGTAGTGGCTGGAAATGAAGCAGGACGAAGGTGGATTC 978
QY 724 CCATGACTAATGGGCAAGCATTTGCTCCTCTGAAAGTTCGGGCGAGTACTGACATCGATG 783
Db 979 CCATGACGATGGCAAGCATTTGCTCCTCTGAGGTCGGGGTGTGGTAGATTTGATG 1038
QY 784 CATCTACTGAATPACACATGGAAGCGCTTTTACTGAATGATGAAACTCGCCAGCCTCTAT 843
Db 1039 CATCAACTGATTACAACTGGAAGATGCTTATTGAAAGCAAGCAAGGCGGAGTTC 1098
QY 844 CTAGAAAAGTCCCATTTGCTTCTCGAATAAATGAAATCCCTAAGATGTCATGTTCTGCG 903
Db 1099 CTAGAAAAGTCCCATTTGCTTCTCGAATAAATCCATACAGGATGTCATTTGCTGTCG 1158
QY 904 GGTTCGTTGTTCTAAGCATTTCTCGTCACTACCGTCTCAAAATCCTGTGCGTAATGCAT 963
Db 1159 GATTGATTTGTTCTAAGCATTTCTTGGCACTACCGTATCAAAATCCTGTGCGCAATGCAT 1218
QY 964 ACCCATGTGGCTTTTATCTGTTATATGTGAGATTTGTTGCTTTATCTCGTACTGCG 1023
Db 1219 ACCCATTTAGCTTCTATCTGTTATATGTGAGATCTGTTGCTCTTTCGTGGATTTGG 1278
QY 1024 ATCAGTTCGCGAAGTGGTTTCCAATCAACCGGGAGACCTTACTTGAATGACGTGCTTAA 1083
Db 1279 ATCAGTTCCTCAAGTGGTTTCCAATCAACCGGGAGACGTTACTTGAATAGGCTGGCATTAA 1338
QY 1084 GGTATGACCGGAGAGGTGAACCGTCTCAGTTGCGTGTGTTGACATATTTGTCAGTACAG 1143
Db 1339 GGTATGACCGGAGAGGTGAGCCATCTCAGTTGCGTGTGTTGACATTTTCGTCAGTACAG 1398
QY 1144 TCGACCCCTTGAAGAGCCACCTATCGTCACTGCCAACACTGTGCTATCCATTTCTGCTG 1203
Db 1399 TCGACCCCAATGAAGGAGCCTCTCTTGTCACTGCCAATAGCGTGTCTATCCATTTCTGCTG 1458
QY 1204 TGGATTTCCGCTGACAAAGTCTCTTGTGATGTATCTGATGACCGGAGCTTCAATGCTGA 1263
Db 1459 TGGATTTACCCCTGTGATGAAGTCTCTTGTGATGTATCTGATGAGGTGCGATGCTGA 1518
QY 1264 CTTTTGACGCTTTGGCTGAGACTTCAGAGTTTGGCTAGGAAATGGGTACCATTTTGTGAAGA 1323
Db 1519 CATTTGATGCACTAGCTGAGACTTCAGAGTTTGGCTAGGAAATGGGTACCATTTTGTGAAGA 1578
QY 1324 AGTATGACATTGAACCCAGAGCTCCCGAGTTTACTTTTGGCAGAAAAATGATTACCTGA 1383
Db 1579 AGTACAACATTGAACCTAGAGCTCTGAATGTAATCTTCTCCAGAAAATGATTACTTGA 1638
QY 1384 AAGACAAAGTCAGGCTTCATTTGTTAAAGCCCGCGGCCATGAAGAGAGAAATATGAAG 1443
Db 1639 AGGACAAAGTCAGGCTTCATTTGTTAAAGCCCGCGGCCATGAAGAGAGAAATATGAAG 1698

QY 1444 AATTTAAATCAGGATAAATGCCCTAGTTTCTAAGSCATTGAAAGTCCCGGAGGAGAT 1503
Db 1699 AATTCAAAGTTAGGTTAAATGGCCCTTGTCTTAAAGCACAGAAAGTTCCTGAGGAAGAT 1758
QY 1504 GGATCATGCAAGATGGCCACCATGGCCAGGAAAAAATACACAGGGATCATCTCTGGAATGA 1563
Db 1759 GGATCATGCAAGATGGCCACCATGGCCAGGAAAAAATACACAGGGACCATCTCTGGAATGA 1818
QY 1564 TTCAGGTTTCTTCCGTTCAGTGTGGCTTGTGATCTAGAGGGTAAATGAGTCCCCCGTT 1623
Db 1819 TTCAGGTTTCTTCCGTTCAGTGTGGCTTGTGATCTAGAGGGCAATGAGTACCCCGTT 1878
QY 1624 TAGTTTATGTGCTCGTGAAGAGCTCTCGGTTCCAGACCCACAGAGAGGCTGTGGCA 1683
Db 1879 TGGTCTATGTTTCTCGTGAAGAGCGTCTCGAATCCAGCATCACAGAAGCTGGTGCCA 1938
QY 1684 TGAATGCCCTTGTTCGTGTCTCAGCTGTCTTACTTAATGGAACAATCATGTTGAATCTTG 1743
Db 1939 TGAATGCTCTTGTTCGTGTCTCAGCTGTCTTACCAATGGAACAATCATGTTGAATCTTG 1998
QY 1744 ATTGTGATCACTACATCAACAAACAGCAAGCGTGTCCGAGAGCTATGTGCTTCCTAATGG 1803
Db 1999 ATTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCCTTATGG 2058
QY 1804 ATCCAAAACCTAGGTCGCGAAGTCTGTATGTGSCAGTCCCAAAAGGTTTGATGCGATTG 1863
Db 2059 ACCCTAACCTAGGAAGGAGTGTCTCTAGCTCAGTTCCTCCAGAGATTCGATGSCATTG 2118
QY 1864 ATAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTTTGATATTAACTTTGAGGGGCC 1923
Db 2119 ACAGCAATGATCGATATGCAACAGGAACACCGTGTCTTCGATATTAACTTTGAGAGGTC 2178
QY 1924 TTGACGCGATTCAAGGACAGTTCATGTGGGAACATGCTGTTGTTTTTCAAGAAACAGCTA 1983
Db 2179 TTGATGGCATCAAGGACAGTTCATGTCCGAACCTGCTGTGTTTCAACCCAGAACAGCTC 2238
QY 1984 TCTATGTTTATGAGCCGCCAAATTAAGGCGAAGAACCCAGGTTTCTTGTGCATCAGTATG 2043
Db 2239 TATATGTTTATGAGCCGCCAAATTAAGCAGAAAGAGGTTGTTTCTTGTCATCAGTATG 2298
QY 2044 GGGCAAGAAGAAGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAAAGTCCAAACA 2103
Db 2299 GCGGTAGGAAGAAGCGCAAGCAAAATCAAGAA---GGGCTCGGACAGAAAGTCCGAGA 2355
QY 2104 AGCATGTGGAAGTCTCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGGTGTGTAAG 2163
Db 2356 AGCATGTGGAAGTCTCTGTCGCAAGTATTCACCTTTGAAGATATAGAGGAGGAGTTGAAG 2415
QY 2164 GTGCTGGGTTTGTAGTACAGAAATCAGTTCATGCTCTCAATGAGCTTAGAGAAAGAT 2223
Db 2416 GCGCTGGATTGTAGCAGCAGAAATCAGTTCATGCTCAAAATGAGCCTGGAGAAAGAT 2475
QY 2224 TTGGCCAGTACAGCAGATTTGTTCCTCCACTCTGATGGAATATGGTGGTGTCTCCTCAGT 2283
Db 2476 TTGGCCAGTCCGCGAGGTTGTTCCTCCACTCTGATGAGTATGTTGGTGTCTCCTCAGT 2535
QY 2284 CCTCCACTCCAGATCTCTTTTGAAGAAAGCTATCCATGTCCATAGTTGTGGCTATGAGG 2343
Db 2536 CCGAACTCCGAGTCTCTTCTGAAAGAAAGCTATCCATGTTTATAAGCTGTGGCTATGAGG 2595
QY 2344 ACAAGTCTGAATGGGAACCTAGATTTGGTGGATCTAGTATCTGTACAGAGATATTC 2403
Db 2596 ACAAGCTGAATGGGAACCTAGATTCGGGTGGATCTACGGTTCTGTGACAGAAACATTC 2655
QY 2404 TTACTGGATTCAAGATGACCGCAAGAGGCTGGCGTTTCACTCTATTGCTATGCCAAGCGCC 2463
Db 2656 TCACGGGATTCAGATGACCGCGGAGGCTGGCGGTGATCTACTGATGCCAAGCGCGC 2715
QY 2464 CAGCTTTCAAGGGATCTGCCGCCCATCAATCTTTTTCAGATCTGTGTAACCAAGTGTGCGGT 2523
Db 2716 CAGCTTTCAAGGGTCTGCCGCCCATCAATCTTTTTCAGACCGCTCTGAACCCAGGTGCTCCGT 2775
QY 2524 GGGCTCTCGGTTCTGTTGAAATTTCTTTTCAGCCGCGCATTTGCCCTTTATGGTATGGCTACG 2583

Db 2776 GGGCTCTGGTCCGTGGAGATCTCTTTCAGCGGCACTGCCCCCTGTGTAGTACGGTACG 2835
Qy 2584 GAGGGCCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCACTAA 2643
Db 2836 GAGGGCCGTCAAGTTCCTGGAGAGATTCGCGTACATCAACACCACTTACCCGCTCA 2895
Qy 2644 CTTCTCTCCGCTTCTAGTCTATTGTATATGCTGCTATCTGTCTGCTCACTGGAAAGT 2703
Db 2896 CGTCCATCCGCTTCTCATCTACTGATCTCTGCGGCCATCTGTCTGCTCACCGAAAGT 2955
Qy 2704 TCATCATGCGAGAGATTAGCAATCTGCGCAGTATCTGGTTCAATTGCGCTCTTCTTTCAA 2763
Db 2956 TCATCATTCAGAGATCAGCAATCTGCGCAGATCTGTTTCACTTCTCTTCACTCGA 3015
Qy 2764 TTTTCCGCACTGATTCCTTAGATGAGGTGGAGTGGTGTGTCATGTAGCAGTGTGGA 2823
Db 3016 TCTTCCGCACTGATTCCTTAGATGAGGTGGAGTGGTGTGTCATGTAGCAGTGTGGA 3075
Qy 2824 GGAATGAACAGTCTCGGTTCATTTGAGGTATCTCTGACATCTGTTGCGCTTCTTCAGG 2883
Db 3076 GGAACGAGAGTCTCGGTTCATTTGAGGTATCTCTGACATCTGTTGCGCTTCTTCAGG 3135
Qy 2884 GTCTTCTGAAGTGTCTGCGGTATCGACACCAATCTTCACTGTCACTCAAAAGGCTAATG 2943
Db 3136 GCTGTCTCAAGTGTCTGCGGTATCGACACCAATCTTCACTGTCACTCAAAAGGCTCGG 3195
Qy 2944 ACAAGAGCGGATCTTGTGAGCTCTACATGTTCAAGTGGAGCAGCTTCTCATCCCTC 3003
Db 3196 ACCAGAGCGGATCTTGTGAGCTCTACATGTTCAAGTGGAGCAGCTTCTCATCCCTC 3255
Qy 3004 CGAGGACCACTTTGATCAATTAACATGTTGGTCTCGTTGCTGCACTCTTACGCCATCA 3063
Db 3256 CCACCACTTCTGATCAATTAACATGTTGGTCTCGTTGCTGCACTCTTACGCCATCA 3315
Qy 3064 ACAGTGTACCAATCATGGGGCCGCTCTTTTGGAGGCTCTTCTTGGCTTCTGGGTGA 3123
Db 3316 ACAGGGATACCACTGTTGGGGCCGCTCTTGGCAAGCTCTTCTCGCTTCTGGGTGA 3375
Qy 3124 TTGTTCACTTATACCAATCTCTCAAGGCTCTTATGGGAGGCAAAACCGGACACACGACGA 3183
Db 3376 TCGTCCACCTGTACCGGCTCTCAAGGGCTCTATGGGAGGCAAAACCGGACACACGACGA 3435
Qy 3184 TTGTCATGCTGGGCTGCTCTCTGCTTCTATCTTCTCTGCTGCTGGGTTCTGTTG 3243
Db 3436 TCGTCTGCTGGGCTCTCTCTGCTTCTATCTTCTCTGCTGCTGGGTTCTGTTG 3495
Qy 3244 ATCCATTCACTACCGGCTCTGCTGGGCCCAATATCCAAACCTGTGGCATCAACTGTAGG 3303
Db 3496 ACCCTTCAACCCCGGCTCACTGGCCGGATACCCAGAGCTGTGGCATCAACTGTAGG 3555
Qy 3304 AAGTGGAGTTT 3316
Db 3556 GAAGTGGAGTT 3568

RESULT 7
US-10-160-719-49
; Sequence 49, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822

; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719-49

Query Match 70.9%; Score 2572.6; DB 15; Length 3746;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

Qy 4 CGAGGAACCCCGTCCAGCTCTGTCTCGTTCGGTTCGGTTCGGATCGCTTCGCGCCGATGG 63
Db 265 CCAGGCCCCAGGCTCCAGGCCAGCTCCCTCGAGCTTCTCGGCGAGCTCGCTTGGCCATGG 324
Qy 64 ACGGCGACGGGACGCCCTGAGTCCGGGAGGCAACGGGCGCGGGGACGTGTGCGCAGATCT 123
Db 325 AGGGCGACGGGACGCCCTGAGTCCGGGAGGCGCGGTGGCGGACAGGTGTGCGCAGATCT 384
Qy 124 GCGCCGACGGCTTGGGACACACAGTTCGAGCGGCGACGCTTTCACCGCTTCGCGAGTCTGCC 183
Db 385 GCGGCGACGGCTTGGGACACACAGTTCGAGCGGCGGAGCGGACGCTTTCGCGGCTTCGCGCTGCG 444
Qy 184 GCTTCCCGTTCGCGCCCTGCTACGAGCAGCGGCAAGAGGGGACCCAGGCTTGC 243
Db 445 GGTTCCTCGTGTGCCGCCCTGCTACGAGTACGAGCGCAAGGACGCGACGCGCGTGC 504
Qy 244 TCAGTGCAGAGCAAGTACAGCGCCACAGAGGGAGCCAGCGATCCGCGGGAGGAAG 303
Db 505 CCAGTGCAGAGCAAGTACAGCGCCACAGGGGAGCCCGCGATCCGTTGGGGAGGAAG 564
Qy 304 GCGACGACACTGATGCGCGATGATGTGTAGTACCTTCAACTACCTTCGCTTCGCGACCTGAGG 363
Db 565 GAGACGACACTGATGCGG-----TAGCGACTTCAATTACCTTGCATCTGGCAATGAGG 618
Qy 364 ACCAGAGCAGAGATGCTGACAGATGCGAGTGGCGGCGATGAACACCGGGGCGAGTG 423
Db 619 ACCAGAGCAGAGATGCTGCGGACAGATGCGAGTGGCGGCGATGAACAGTTTGGGGGCGAG 678
Qy 424 GCAATGCTGGCCACCCCAAGTATGACAGTGGCGGAGATCGGCCCTCTCCAAGTATGACAGTG 483
Db 679 GGGATGTTGTCGCCCGCAAGTATGACAGTGGCGGAGATCGGGCTTACCAAGTATGACAGTG 738
Qy 484 GAGAGATCCCTAGGGGATACGTCCTTCAGTCAACCAACAGCCAGATGTTCAGGAGAAATCC 543
Db 739 GCGAGATTCCTCGGGGATACATCCCATCAGTCACTAACAGCCAGATCTCAGGAGAAATCC 798
Qy 544 CTGGAGCTTCGCTGATCATCATGATGTCCTTACGAGGAAACATCAGCAGACGTCGCTC 603
Db 799 CTGGTGTTCCTCGGCGATCATATGATGTATGATGTATGATGTATGATGTATGATGTATGATG 858
Qy 604 CGTTTCCCTATGTAATCATTCACCAATCCGTCAGGGGAGTTCCTCGGCGAGTATTGGGA 663
Db 859 CATTTCCCTATGTAATCATTCGCGCAATCCGTCAGGGGAGTTCCTCGGCGAGTATTGGGA 918
Qy 664 ATGTTGCTGGAAAGAGAGATGATGGCTGGAAAAATGAAGCAGGCAAGGGGTGCGATTC 723
Db 919 ATGTTGCTGGAAAGAGAGAGTGTGATGGCTGGAAAAATGAAGCAGGCAAGGGGACGATTC 978
Qy 724 CATGACTATGGCAAGCATGCTCTCTGAGGTTCGCGGAGTTCGAGCTACTGACATCGATG 783
Db 979 CCATGCAATGGCAAGCATGCTCTCTGAGGTTCGCGGAGTTCGAGCTACTGACATCGATG 1038
Qy 784 CATCTACTGAATACAAACATGGAAGAGCGCTTTCATGTAATGATGAAGTTCGCGAGCTCTAT 843
Db 1039 CATCACTGATTAACATGGAAGAGTGCCTTATGTAACAGCAAGTTCGAGCGCTCTAT 1098
Qy 844 CTAGAAAGTCCCATTTGCTTCTCCAAATAAATCCCTACAGAAATGCTTCTGTTCTGCG 903

Db 1099 CTAGGAAGTTCCTCTCCAGGATAAATCCATACAGGATGGTCAATGTGCTGC 1158
QY 904 GGTGTTGTTCTAAGCATCTTCCTGCACACTACCGTCTCAAAATCCTGTGCGTAATGAT 963
Db 1159 GATTGATGTTCTAAGCATCTTCCTGCACACTACCGTCTCAAAATCCTGTGCGCAATGCAAT 1218
QY 964 ACCCACTGTGGCTTTATCTGTATATGTAGATTTGGTTTCTTTATCTCGTACTGG 1023
Db 1219 ACCCATATGGCTTCTATCTGTATATGTAGATCTGGTTTCTTTCTGTGGATATGG 1278
QY 1024 ATCACTCCGGAAGTGGTTTCCAAATCAACCGGGAGACCTACCTTGATAGACTGGCTTTAA 1083
Db 1279 ATCACTCCGGAAGTGGTTTCCAAATCAACCGGGAGAGGTACCTTGATAGGCTGGCATTA 1338
QY 1084 GGTATGACCGAAGAGGTGAACCGTCTCAGTTGGCTGTGTGACATATTTGTCACTACAG 1143
Db 1339 GGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGTGTGACATATTTGTCACTACAG 1398
QY 1144 TCGACCCCTTGAAGAGGCCACTATCGTCACTGCCAACACTGTGCTATCCATCTTGTG 1203
Db 1399 TCGACCCCAATGAGAGGCTCTCTGTGCTACCTGCGCAATACCGTGTCTATCCATCTTGTG 1458
QY 1204 TTGATATCCCGTGACAAAGGTCTCTTGTCTATGTATCTGATGACGAGGCTTCAATGCTGA 1263
Db 1459 TGGATTACCTGTGATAAGGTCTCTTGTCTATGTATCTGATGAGTGGTGGATGCTGA 1518
QY 1264 CTTTGTAGCGCATGCTGAGACTTCAGAGTTTGTCTAGGAATGGGTACATTTGTGAAGA 1323
Db 1519 CATTTGATGCACTAGCTGAGACTTCAGAGTTTGTCTAGAAAATGGGTACCATTTGTTAAGA 1578
QY 1324 AGTATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTGGCCAGAAAATGATTAACCTGA 1383
Db 1579 AGTACAAATTTGAACCTAGAGCTCTGATGGTACTTCTCCAGAAAATGATTAACCTGA 1638
QY 1384 AAGCAAAAGTCCAGCTTCAATTTGTTAAAGACCGCCGGCCCATGAAGAGANAATGAAG 1443
Db 1639 AGGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCCGGCCCATGAAGAGANAATGAAG 1698
QY 1444 AATTAAATCAGNATAAATGCCCTAGTTTCTAAGGCATTTGAAGTCCCGAGGAAGAT 1503
Db 1699 AATTCAAAGTTAGGGTAAATGGCTTGTGTGTAAGGCACAGAAAGTTCTCGAGGAAGAT 1758
QY 1504 GGATCATCAAGATGGCACACATCGGCCAGGAACAATACAGGAGATCATCTGGAATGA 1563
Db 1759 GGATCATGCRAGATGGCACACATGGCCAGGAACAATACAGGAGCATCTGGAATGA 1818
QY 1564 TTCAGGTTTTCTTGGTCAAGTGGTGGCTTGTATCTGATGAGGTAATGAGCTCCCGCGTT 1623
Db 1819 TTCAGGTTTTCTTGGTCAAGTGGTGGCTTGTATCTGATGAGGCAATGAGCTACCCCGTT 1878
QY 1624 TAGTTTATGTCTCGTGAAAGCGTCTGGTTCCAGCACACAGAGGCTGTGGCCA 1683
Db 1879 TGGTCTATGTTCTCGTGAAAGCGTCTGGATTCAGATTCAGAGATCAACAAGAAAGTGGTCCA 1938
QY 1684 TGAATGCCCTTGTTCGTCTCAGCTGTCCCTTACTAATGGAACAATACATGTTGATCTTG 1743
Db 1939 TGAATGCTCTTGTTCGTCTCAGCTGTGCTTACATGGAACAATACATGTTGATCTTG 1998
QY 1744 ATTGTGATCACTACATCAACAACAGCAGGCTGTCCGAGAGACTATGTGCTTCCCTAATGG 1803
Db 1999 ATTGTGATCACTACATTAACAACAGTAAAGCTCTCAGGAGACTATGTGCTTCCCTAATGG 2058
QY 1804 ATCCAAACCTAGGTCGCGAAGTCTGTATGTGAGTGTCCCAAAAGGTTTGAATGGGATG 1863
Db 2059 ACCCTAACCTAGGAAGGAGTGTCTGTCTACGTCCAGATTTCCCGACAGAGATTCGATGGCAT 2118
QY 1864 ATAGAATGATCGATATGCAAAACAGGAACAATGCTTTTGTATATTAACCTTGAAGGCC 1923
Db 2119 ACAGGAATGATCGATATGCCAAACAGGAACAACCGTGTCTTTCGATATTAACCTTGAAGG 2178
QY 1924 TTGACCGCAATTAAGGACCACTTTATGTGGGAATGGTTGTGTTTCAACAGAACAGCTA 1983

Db 2179 TTGATGGCATCAAGGACCAAGTTTATGTATCGGAACCTGGCTGTGTTTCAACGAAACAGCTC 2238
QY 1984 TCTATGTTATGAGCCCCCAATTAAGCGGAAAGAACCAAGGTTTCTTGGCATCACTATGTG 2043
Db 2239 TATATGTTATGAGCCCCCAATTAAGCAGAAAGAGGTTGTTTCTTGTCTCATCACTATGTG 2298
QY 2044 GGGGCAAGAAGAACCAAGCAAGTCAAAAGAAAGAGGCTCAGATAGAAAAGTTCGAACA 2103
Db 2299 GCGGTAGGAAGAACCAAGCAAGCAAAATCAAGAA---GGGCTCGGACAAAGAAAGTCCGAGA 2355
QY 2104 AGCATGTGACAGATTTCTGTTCAGTATTCATCTCGAAGACATAGAGAGGTTGTGAAG 2163
Db 2356 AGCATGTGACAGATTTCTGTGCCAGTATTCACCTTGAAGATATAGAGAGGAGTTGAAG 2415
QY 2164 GTGCTGGGTTTGAATGATGAGAAATCAGTTCTCATGCTCAAAATGAGCTTGAAGAAAGAT 2223
Db 2416 GCGCTGGATTTGACGACGAGAAATCACTTCTTATGTCTCAAAATGAGCCTGGAGAAGAT 2475
QY 2224 TTGGCCAGTCAAGCAGATTTGCTGCCTCCACTCTGATGGAATATGGTGGTGTCTCAGT 2283
Db 2476 TTGGCCAGTCCGACAGGTTGTGTGCTCCACTCTGATGAGATATGGTGGTGTCTCAGT 2535
QY 2284 COTCACTCCAGAAATCTCTTTTGAAGAAAGCTATCCATGTCAATAAGTTGTGGCTATGAG 2343
Db 2536 CGCAACTCCGAGTCTCTTCTGAAAGAGCTATCCATGTTATAGCTGTGGCTATGAG 2595
QY 2344 ACAAGTCTGAATGGGAACTCAGATTTGGTTGGATCTATGGATCTGTCTCAGAGATATTC 2403
Db 2596 ACAAGACTGAATGGGAACTCAGATCGGTTGGATCTACGGTCTCTGTGACAGAAACATTC 2655
QY 2404 TTACTGGATTCACAGATGACACGACAGAGCTGGCGTTTCAAGATCTGTGAAACAGCTGCTG 2463
Db 2656 TCACCGGATTCACAGATGACACGCGGAGGCTGGCGTGGATCTACTGCAATGCCCAAGCGC 2715
QY 2464 CAGCTTTCAAGGGATCTGCCGCCATCAATCTTTCAAGATCTGTGAAACAGCTGCTGCGGT 2523
Db 2716 CAGCTTTCAAGGGTCTGCCGCCATCAATCTTTGGACCGTCTGNAACAGGTTGCTCCGT 2775
QY 2524 GGGCTCTCGTCTGTGTTGAAATCTTTTACGCGGCAATGCCCTTATGGTATGCTACG 2583
Db 2776 GGGCTCTGGGTCCGTGGAGATCTCTTTCAGCGGCACTGCCCTGTGGTACGSCATG 2835
QY 2584 GAGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACACCACTTTACCCACTAA 2643
Db 2836 GAGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCACTTTACCCCGTCA 2895
QY 2644 CCTCTCTCCCGCTTCTAGTCTATTTGATATTTAGCTGCTATCTGTCTGCTCACTGGAAGT 2703
Db 2896 GGTCCATCCCGCTTCTCATCTACTGCTCTGCGCGCATCTGTCTGCTACCGGAAAGT 2955
QY 2704 TCATCATGCGAGATTTAGCAACTTGGCCAGTATCTGGTTCAITGGCTCTTCCCTTCAA 2763
Db 2956 TCATCATTTCCAGATCAGCAACTTTCGCCAGCATCTGGTTTCAITCTCCCTCTTCACTCGA 3015
QY 2764 TTTTTCGCCACTGTGATCTTGTAGATGAGTGGTGGTGTGGCATTTGAGCATTTGAGGTTGGA 2823
Db 3016 TCTTCGCCACGCGCATCTCTGGAGATGAGTGGAGCGGGGTGGCATTCGACGAGTGGTGA 3075
QY 2824 GGAATGAACAGTCTCTGGTCAITGGAGTATCTCTGCACATCTGTTTGGCGCTTCTTCAAG 2883
Db 3076 GGAACGAGCAGTCTCTGGTGTCTGGGGCATCTCCGCGCATCTTTCGCGCTGTTCGAG 3135
QY 2884 GTCTTCTGAAGGTGCTTTCGCGGTATCGACCACTTCACTGTCACTCACTCAAGGCTAATG 2943
Db 3136 GCCTGTCTCAAGGTGCTGCGCGCATCGACCACTTTCACCGTCACTCAAGGCTCGG 3195
QY 2944 ACNAGAGGCGGCTTTCGCTGAGCTTACATGTTCAAGTGAACGAGCTTCTCATCCCTC 3003
Db 3196 ACGAGGACGCGAGCTTCGCGAGCTGTACATGTTTCAAGTGAACGAGCTCTCTGATCCCGC 3255
QY 3004 CGACGACCATTTTGTATTAACATGTTGTGTTGTGCTGGCACCTCTTACGAGCTCA 3063
Db 3256 CCACCACTCTGATCATCAACTGGTCTGGCGCTGCTGGCGCATCTCTTACGCCATCA 3315

QY 1384 AAGACAAAGTCAGCCTTCATTTGTTAAAGACGCGCGGCCCATGAAGAGAGAAATATGAAG 1443
DB 1656 AGGACAAAGTCAGCCTTCATTTGTTAAAGACGCGCGGCCCATGAAGAGAGAAATATGAAG 1715
QY 1444 AATTTAAATACAGGATAAATGCGCTAGTTTCTAAGSCATTTGAAGTCCCGGAGGAGGAT 1503
DB 1716 AATTCAAAGTTAGGGTAAATGCGCTAGTTTCTAAGSCACAGAAAGTTCTCTGAGGAAGAT 1775
QY 1504 GGATCATGAGATGGCACACCATGGCAGGAGAAACATACAGGAGATCATCTGGAATGA 1563
DB 1776 GGATCATGAGATGGCACACCATGGCAGGAGAAACATACAGGAGATCATCTGGAATGA 1835
QY 1564 TTACAGTTTCTCGTGGTCAAGTCAGTGTGGCTTTGATACCTGAGGTAAATGAGTCCCGGTT 1623
DB 1836 TTCAGTTTCTCGTGGTCAAGTCAGTGTGGCTTTGATACCTGAGGCAATGAGTACCCGTT 1895
QY 1624 TAGTTTATGTGTCCTGTAAGAACGCTCTGGGTTCCAGCACACAGAGAGGCTGGTGCCA 1683
DB 1896 TGGTCTATGTTTCTCGTGAAGAACGCTCTGGATTCAGCATCACAAAGAAAGCTGGTGCCA 1955
QY 1684 TGAATGCCCTTGTTCGTGTCAGCTGCTCTTACTTAATGGAACAATACATGTTGAATCTTG 1743
DB 1956 TGAATGCTCTTGTTCGTGTCAGCTGCTTACCAATGGACAATACATGTTGAATCTTG 2015
QY 1744 ATTGTGATCACTACATCAACACAGCAAGGCTGTCCGAGAGCTATGTCTTCTTAATGG 1803
DB 2016 ATTGTGATCACTACATTAACACAGTAAGGCTCTCAGGAGCTATGTCTTCTTAATGG 2075
QY 1804 ATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAGTTCGCCAAAGGTTTGTAGGATGG 1863
DB 2076 ACCCTAAGCTAGGAAGAGTGTCTGCTAGCTCCAGTTCCTCCAGAGATTTCGATGCAATTG 2135
QY 1864 ATAGGAATGATCGATATCAACACAGAACGCTGTCCGAGAGCTATGTCTTCTTAACAGCA 1923
DB 2136 ACAGGAATGATCGATATGCAACAGGAACACCGTGTTCCTGATATTAACCTTGAAGGTC 2195
QY 1924 TTGACGGCATTCAGGACACAGTTTATGTGGAACTGGTGTCTTTTCAACAGAACAGCTA 1983
DB 2196 TTGATGGCATCCAGGACACAGTTTATGTGCGNACTGGCTGTGTTTCAACCGAACAGCTC 2255
QY 1984 TCTATGTTATGAGCCCAATTAAGGCGAAGAACCCAGGTTTCTTGGCATCACTATGTTG 2043
DB 2256 TATATGTTATGAGCCCAATTAAGCAGAGAGAGGTTGTTTCTGTCTCATCACTATGTTG 2315
QY 2044 GGGGCAAGAGAGGACGACGAGTCAAGAGAGAGAGCTCAGATGAAGAAAGTTCGAACA 2103
DB 2316 GCGGTAGGAAGAGCAAGCAAAATCAAGAA ---GGGCTCGGACAGAGAAAGTTCGAGA 2372
QY 2104 AGCATGTGACAGATTCCTGTCAGTATTCATCTCGAAGACATAGAGAGGAGTGTGAAG 2163
DB 2373 AGCATGTGACAGATTCCTGTCAGTATTCATCTTGAAGATATAGAGAGGAGTGTGAAG 2432
QY 2164 GTGCTGGTTTGAATGAGAAATCAGTTCTCATGTCCTCAATGAGCTTAGAGAGAGAT 2223
DB 2433 GCGCTGGAATTCAGCACGAGAAATCACTTCTTATGTCCTCAATGAGCTCGAGAGAGAT 2492
QY 2224 TTGSCAGTACAGCAATTTGTGCTCCACTCTGATGGAATATGGTGGTTCCTCAGT 2283
DB 2493 TTGSCAGTCCGAGCGTTTGTGCTCCACTCTGATGGAATATGGTGGTTCCTCAGT 2552
QY 2284 CTTCACTCCAGAAATCTCTTTTGAAGAGAGCTATCCATGTATAGTTTGGCTATGAGG 2343
DB 2553 CGCAACTCCGAGTCTCTCTGAAAGAGCTATCCATGTTATAGCTGTGGCTATGAGG 2612
QY 2344 ACAAGTCTGAATGGGAACCTGAGATTTGGATTTGATCTATGATCTGTCTCAGAGAAATTC 2403
DB 2613 ACAAGACTGAATGGGAACCTGAGATTCGGGTGGATCTACGTTCTGTGTGACAGAGACATTC 2672
QY 2404 TTACTGGATTCAGATTCAGCCAGAGAGGCTGGCTTCACTCTATGATGATGCTCCAGAGGCC 2463
DB 2673 TCACCGGATTCAGATTCAGCCAGAGAGGCTGGCTGGCTGATCTACTGATGCTCCAGAGGCC 2732
QY 2464 CAGCTTTCAAGGATTCGCCCCCATCAATCTTTTCAGATCGTCTGTAACCAAGTGTGCGGT 2523

DB 2733 CAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTGGAACCGTCTGAACACAGGTGCTCCGTT 2792
QY 2524 GGGCTCTCGTCTCTGTGAATTTCTTTTTCAGCGCGCATTTGCCCTTATGGTATGGTACG 2583
DB 2793 GGGCTCTTGGGTCCGTTGGAGATCTCTTTCAGCGCGCATTTGCCCTTGTGGTACGGTACG 2852
QY 2584 GAGGGGCGCTCAAGTTCTCTGGAGAGATTGCGTTTATCAACAACCAACCATTTACCACTAA 2643
DB 2853 GAGGGCGGCTCAAGTTCTCTGGAGAGATTGCGTTTATCAACAACCAACCATTTACCTCA 2912
QY 2644 CTTCTCTCCGCTCTCTAGTCTATTTATATTGCGTCTATCTGTGCTCACTGGAAGT 2703
DB 2913 GGTTCATCCGCTCTCTATCTACTGCACTCTGCCCGCATCTGTGTCTCAACCGGAAT 2972
QY 2704 TCATCATGCCAGAGATTAGCAACTTTGGCAGTATCTGTGTTCAATTCGCTCTTCTTTCAA 2763
DB 2973 TCATCATTCAGAGATCAGCAACTTTGCCAGCATCTGGTTCATCTCCCTCTTCACTCGA 3032
QY 2764 TTTTCGCCACTGGTATCTTGGAGTGAAGTGGTGGTATTTGCAATTGACAGTGGTGA 2823
DB 3033 TCTTCGCCACGGGCATCTCTGGAGATGAGTGGAGCGGGTGGGCATTCGACGAGTGGTGA 3092
QY 2824 GGAATGAACAGTTCTGGGTCAATTGGAGTATCTCTGCACTCTGTGTCGCTCTTCAGG 2883
DB 3093 GGAACGACAGTTCTGGGTGATCGGGGCACTCTCCGCGACCTCTTCGCGGTGTTCCAG 3152
QY 2884 GTCTTCTCAAGTGTCTTCCGATTCGACACAACTTCACTGTCACTCAAAAGGCTAATG 2943
DB 3153 GCTGCTCAAGTGTCTGCGCGCATCGACACAACTTCACTGTCACTTCAAGGCTCGG 3212
QY 2944 AGGAAGAGGCGACTTTCGAGCTCTACATGTTCAAGTGGACAGCGTTCATCTCCCTC 3003
DB 3213 ACAGGACGCGACTTTCGAGCTCTACATGTTCAAGTGGACAGCGTTCATCTCCCTC 3272
QY 3004 CGACGACCATTTTGATCAATTAACATGTTGGTGTGTTGCTGGCACCTCTACGCACTCA 3063
DB 3273 COACACCATCTCTGATCAATCAACCTGGTGGGTCTGTCGCGGCACTCTCTACGCACTCA 3332
QY 3064 ACAGTGTTCACCAATCATGGGGCGGCTCTTTGGGAAGCTCTTTTGGCTTCTGGGTGA 3123
DB 3333 ACAGCGGATACAGTCTGGTGGGCGGCTCTTCGGAAGCTCTTCTGGCTTCTGGGTCA 3392
QY 3124 TTGTTCACTTATCCATTTCTCAAGGTCTTATGGGAGGCAAAACCGCACACCGAGCA 3183
DB 3393 TCGTCCACTGTACCGTTCTCTCAAGGCTCTCATGGGCGCTCATGGGAGGAGCAACCGCA 3452
QY 3184 TTGTCATGCTGGGCT 3243
DB 3453 TCGTCTGCTGGGCT 3512
QY 3244 ATCCATTCACCTACCGCT 3303
DB 3513 ACCCTTCAACACCGGCT 3572
QY 3304 AAAGTGGAGTTT 3316
DB 3573 AAAGTGGAGTTT 3585

RESULT 9

US-10-209-059-9
; Sequence 9, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822

```
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-9

Query Match      70.9%; Score 2572.6; DB 15; Length 3773;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

QY      4 CGAGGAACCCCGCTCCAGCTCTCTCGTGGTGGGCTTGATCGCTCTGCGCGCCATGG 63
DB      282 CAGGGCCCAAGCTCCAGGCCAGCTCCCTCGACGTTTCTCGGCGAGCTCGCTTGCCATGG 341

QY      64 ACGGCGACGCGGACGCCCTGAAGTCCGGGAGGACACGGGGCGGGGAGGAGTGCCAGATCT 123
DB      342 AGGGCGACGCGGACGGCGTGAAGTCCGGGAGGCGGCTGGCGGACAGGTTGCGCAGATCT 401

QY      124 GCGCGACGCGCTGGGACCAACCTTGGACGGCGACGCTCTTCAACGCTCGACGCTCTGCC 183
DB      402 GCGGCGACGCGCTGGGACCAACGCGGCGAGGGGACGCTTTCGCGCGCTCGGACGCTCGG 461

QY      184 GCTTCCGCGCTCGCGCCCTGCTACGACGACGAGCGGCAAGGAGGCGACCCAGGCGCTGCC 243
DB      462 GGTTCGCGTGTGCGGCGCTCTGCTACGAGTACGAGCGGAGGACGCGACGCGGCTGCC 521

QY      244 TCAGTGCAAGCAAGTACAGCGCCACAGAGGAGCCAGGCGATCCGCGGGGAGGAAG 303
DB      522 CCCAGTGCAAGCAAGTACAGCGCCACAGAGGAGGCGCGGCGATCCGTTGGGGGAGGAAG 581

QY      304 GCGACGACCTGATGCGGATGATGCTAGTACCTTCACTACCTGCTGCTGCGACCTGAGG 363
DB      582 GAGACGACCTGATGCGGATGATGCTAGTACCTTCACTACCTGCTGCTGCGACCTGAGG 635

QY      364 ACCAGAGCAGAGATGCTGACAGGATGCGGAGTGGCGCATGAACACCGGGGCGAGTG 423
DB      636 ACCAGAGCAGAGATGCTGCGGACAGATGCGGAGTGGCGCATGAACGTTGGGGGCGAGG 695

QY      424 GCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAAGTATGACAGTG 483
DB      696 GGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGCTTACCAAGTATGACAGTG 755

QY      484 GAGAGATCCCTAGGGGATAGTCCCTTCAGTCAACACAGCCAGATGTCAGGAGAAATCC 543
DB      756 GCGAGATTCCTCGGGGATACATCCCATCAGTCACTACAGCCAGATCTCAGGAGAAATCC 815

QY      544 CTGGAGCTTCGCTGATCATCATGATGTCCTCTACGGGAAACATCAGCAGAGCTGCTC 603
DB      816 CTGGTCTTCCCTGACCATCATGATGATGTCCTCAACTGGGAAATTTGGCAAGCGTCTC 875

QY      604 CGTTTCCCTATGGAATCATTCACCAATCCGTTCAAGGAGTTCCTCGGCGAGTATGGGA 663
DB      876 CATTTCCCTATGGAACCATTCGCGCAATCCGTTCAAGGAGTTCCTCGGCGAGTATGGGA 935

QY      664 ATGTTGCTCGGAAGAGAGAGTTCATGCTGCGGAAATGAAGCAGGCAAGGGTGCAGTTC 723
DB      936 ATGTTGCTCGGAAGAGAGAGTTCATGCTGCGGAAATGAAGCAGGCAAGGGGAGGATTC 995

QY      724 CCATGACTAATGGGACAGCATTCCTCCCTCTGAAGTTCGGGAGCTACTGACATCGATG 783
DB      996 CCATGAGCAATGGCAGCAAGCATTCCTCCCTCTGAGGGTTCGGGTGTTGGTCAATTCATG 1055

QY      784 CATCTACTGATATCAACATGGAAGACCTTTACTGAATGATGAACCTCGCCAGCGCTCTAT 843
DB      1056 CATCACTGATATCAACATGGAAGATGCTTTATTTGAACGACGAAATTCGACAGCGCTCTAT 1115
```

```
QY      844 CTAGAAAAGTCCCCCATTTGCTTCTCCAAAATAAAATCCCTACAGAAATGGTCATTGTCTGC 903
DB      1116 CTAGGAAAGTCCCACTTCTCTCCAGGATAAAATCCATACAGGATGGTCATTGTGCTGC 1175

QY      904 GGTGGTTGTTTAAGCATCTTCTGCACTACGGTCTCAAAATCCTGTGCGTAATGCAT 963
DB      1176 GATTGATTGTTCTAAGCATCTTCTTGCACCTACCGTATCAAAATCCTGTGCGCAATGCAT 1235

QY      964 ACCCACTGTGGCTTTTATCTGTTATATGTAGATTGGTTTGTCTTATCCTGATACTGG 1023
DB      1236 ACCCATTTGGCTTCTATCTGTTATATGTAGATCTGGTTTGTCTTCTGTTGATTTGG 1295

QY      1024 ATCAGTTCGCCGAAGTGGTTTTCAATCAACCGGAGACCTACCTTGTATAGACTCGCTTTAA 1083
DB      1296 ATCAGTTCCTTAAGTGGTTTTCAATCAACCGGAGACGTACCTTGTATAGCTGGCATTA 1355

QY      1084 GGTATGACGAGAGGTGAACCGTCTCAGTTGGTCTGTGTGACATATTTGTGACGTACAG 1143
DB      1356 GGTATGACGAGAGGTGAGCCATCTCAGTTGGTCTGTGTGACATTTTGTGTCAGTACAG 1415

QY      1144 TCGACCCCTTGAAGGAGGCCACCTATCGTCACTGCCAACACTGTGCTATCCATTCTTGTG 1203
DB      1416 TCGACCAATGAAGGAGCCTCCTCTTGTCACTGCCAATACCGTGTCTATCCATTCTTGTG 1475

QY      1204 TTGATTATCCCGTGGCAAGGCTCTTGTCTATGTATCTGATGACGAGCTTCAATGCTGA 1263
DB      1476 TGGATTACCTCTGGGATAAGGCTCTTGTCTATGTATCTGATGAGGCTGCGATGCTGA 1535

QY      1264 CTTTTCAGCATTTGGCTGAGACTTCAAGTCTTCTAGGAAATGGGTACCAATTTGTGAGA 1323
DB      1536 CATTTGATGACCTAGCTGAGACTTCAAGTCTTCTAGGAAATGGGTACCAATTTGTGAGA 1595

QY      1324 AGTATGACATTTGAACCCAGAGCTCCCGAGTTTCTTCTTCCAGAAAATTTGATTACCTGA 1383
DB      1596 AGTATGACATTTGAACCCAGAGCTCCCGAGTTTCTTCTTCCAGAAAATTTGATTACCTGA 1655

QY      1384 AAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGATATGAAG 1443
DB      1656 AGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGATATGAAG 1715

QY      1444 AATTTAAATCAGGATTAATCCCTAGTTTCTAAGGCATTTGAAGTCCCGAGGAGGAT 1503
DB      1716 AATTTAAAGTATGGGTAAATGGCTTGTGCTAAGGCACAGAAAGTTCTCTGAGGAGGAT 1775

QY      1504 GGATCATGCAAGATGCGACACCATGCGCGGAAACAAATACAGGGATCATCTCTGGAATGA 1563
DB      1776 GGATCATGCAAGATGCGACACCATGCGCGGAAACAAATACAGGGACCACTCTCTGGAATGA 1835

QY      1564 TTAGGTTTCTTGGTCAAGTGGTGGCTTGTATCTAGAGGTATAGAGTCCCGCTT 1623
DB      1836 TTAGGTTTCTTGGTCAAGTGGTGGCTTGTATCTAGAGGTATAGAGTCCCGCTT 1895

QY      1624 TAGTTTATGCTCTCGTGAAGAGCTCTCGGTTTCCAGACCCACAGAAAGCTGGTGCCA 1683
DB      1896 TGGTCTATGTTCTCGTGAAGAGCTCTCGGTTTCCAGATTCACAGAAAGCTGGTGCCA 1955

QY      1684 TGAATGCCCTTGTGTTGCTGCTCAGCTGCTCTTACTAATGGACAAATACATGTTGAATCTTG 1743
DB      1956 TGAATGCTCTTGTGTTGCTGCTCAGCTGCTTACCAATGCAATAATCATGTTGAATCTTG 2015

QY      1744 ATTGTGATCATCATCAACAGCAGGCTGTCCGAGAGCTATGCTTCTCTAATGG 1803
DB      2016 ATTGTGATCATCATCAACAGCAGTAAAGCTCTCAGGGAAGCTATGCTTCTCTAATGG 2075

QY      1804 ATCCAACTAGTTCGCGAGTCTGTTATGTCAGTTTCCACAAAGGTTTGTATGGGATG 1863
DB      2076 ACCCTAACCTAGGAGAGGATGCTGCTAAGTCCAGTTTCCAGAGATTCAGTGGCATTG 2135

QY      1864 ATAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTGTATTAATTAATTTGAGGGGCC 1923
DB      2136 ACAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTGTATTAATTAATTTGAGGGGCC 2195

QY      1924 TTGACGGCATTCAGAGGACGAGTTTATGCTGGAACTGGTTGTTTCAACAGAACAGCTA 1983
```



```
Db 2196 TTGATGGCATCAAGGACAGTTTATGTCGGAACCTGGCTGTGTTTCAACCGAAGAGCTC 2255
QY 1984 TCTATGGTTATAGCCGCCCAATTAAGCGCAAGAGCCAGGTTCTTGGCAATCACTATGTG 2043
Db 2256 TATATGGTTATAGCCGCCCAATTAAGCGCAAGAGGAGGTTGTTCTTGTCTATCACTATGTG 2315
QY 2044 GGGCAAGAAGAAGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAGAAAGAGTCAACCA 2103
Db 2316 GGGTAGAGAGAGCAAGCAAAATCAAGAA--GGGCTCGGACAGAAAGAGTCGCAGA 2372
QY 2104 AGCATGTGACAGTTCTGTTCAGTATTCATATCTCGAAGACATAGAGGAGGTTGTGAAG 2163
Db 2373 AGCATGTGACAGTTCTGTGCCAGTATTCACACTTGAAGATATAGAGGAGGTTGAAG 2432
QY 2164 GTGCTGGGTTTGAATGAGAAATCAGTCTCATGTCTCAATAGAGCTTGAAGAGAGAT 2223
Db 2433 GGGCTGGGATTTGACACGAGAAATCACTTCTATGTCTCAATGAGCCTGGAGAGAGAT 2492
QY 2224 TTGGCAGTTCAGCAGCATTTGTTGCTCCACTCTCATGGAATATGTTGTTCTCAGT 2283
Db 2493 TTGGCAGTTCGACAGGTTTGTGCTCCACTCTCATGAGGATGTTGGTGTCTCAGT 2552
QY 2284 CTTCACTCCAGATCTCTTTGAAGAAGCTATCCATGTCTATAAGTTGTGGCTATGAG 2343
Db 2553 CGCAACTCCGAGTCTCTCTGAAGAAGCTATCCATGTTATAAGCTGTGCTATGAG 2612
QY 2344 ACAAGTCTGAATGGGAATGAGATGGTTGATCTATGGATCTGTCAAGAGATATTC 2403
Db 2613 ACAAGACTGAATGGGGAATGAGATCGGTTGATCTACGGTTCTGTGACAGAGACATTC 2672
QY 2404 TTACTGGATTCAGATGACGACGAGGCTGCGTTTCACTATGTCATGTCGCCACAGCGCC 2463
Db 2673 TCACCGGATTCAGATGACGCGCGAGGCTGGGTCGATCTACTGCAATGCCCAAGCGGC 2732
QY 2464 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTAGATCGTCTGAACCAAGTGTGCGGT 2523
Db 2733 CAGCTTTCAAGGGTCTGCCCCCATCANCTTTGCGACGGTCTGAACCAAGTGTCCGGT 2792
QY 2524 GGGCTCTGGTCTGTTGAATCTTTTACGCGGCATGCCCCCTATGTTGGTATGCTACG 2583
Db 2793 GGGCTCTGGTCTGGTGGAGATCTCTTTCAGCGGCACGCCCCCTGTGTTACGCTACG 2852
QY 2584 GAGGCGGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACACCACTTTACCCACTAA 2643
Db 2853 GAGGCGGCTCAAGTTCCTGGAGAGATTCGCGTATCAACACCACTATACCGGCTCA 2912
QY 2644 CCTCTCTCCGCTTCTAGTCTATTGTATATTGCGTCTATCTGTCTCACTGGAAGT 2703
Db 2913 CGTCCATCCGCTTCTCATCTACTGATCTCTGCGGCCATCTGTCTGCTCACCGGAAGT 2972
QY 2704 TCATATGCCAGAGATTAGCAACTTGGCCAGTATCTGGCTCAATGCGGCTCTTCTTTCAA 2763
Db 2973 TCATATTCAGAGATCAGCAACTTCGCGAGCATCTGGTTTCACTCTCCCTCTTCTCTGA 3032
QY 2764 TTTTCGCCACTGGTATCCTTCAGATAGGTCGAGTGGTGGCATTCAGATTCAGGAGTGA 2823
Db 3033 TCTTCGCCACGGGCATCTCTGGAGATGAGTGGAGCGGGTGGGCATTCAGCAGTGGTGA 3092
QY 2824 GGAATGAACAGTCTTGGGTCAATGGAGGTATCTCTGCAATCTGTTTTCGCGCTCTTTCAAG 2883
Db 3093 GGAACGAGCAGTCTCTGGTGTATCGGGGCATCTCCGCGCACTCTTTCGCGCTGTTCAGG 3152
QY 2884 GTCTTCTGAAGTCTTCCCGGTATCGACACAATCTTCACTGTCACTTCAAGGCTAATG 2943
Db 3153 GCCTGCTCAAGTGTGCGCGGCATTCGACACCAACTTCACCGTCACTTCCCAAGGCTCGG 3212
QY 2944 ACGAAGAGGAGCTTTGCTCAGCTCTACATGTTCAAGTGAACGACGCTTCTATCCCTC 3003
Db 3213 ACGAGGACGGGACTTTCGCGAGCTGTACATGTTCAAGTGAACGACGCTCTCTGATCCCG 3272
QY 3004 CGACGACCATTTGATCATTAACATGTTGTGTTGCTGTGCGACCTCTTACGCCATCA 3063
```

```
Db 3273 CCACCACCATCTCTGATCATCAACCTGTCGGCGTCTGTCGCCGCAATCTCTCTAGCCATCA 3332
QY 3064 ACAGTGGTTACCAATCATGGGGCGGCTCTTTTGGAAAGCTCTTTTGGCTTCTTGGGTGA 3123
Db 3333 ACAGCGGATACAGTCTGTTGGGCGCGCTCTTTCGGCAAGCTCTTCTTCCGCTTGGGTGA 3392
QY 3124 TTGTTCACTTATACCATCTCTCAAGGGTCTTATGGGCAAGCAAAACCGCACACCGACGA 3183
Db 3393 TCGTCCACTGTCACCGTTCTCTCAAGGGCTCTATGGGAGGAGAACCGCACCCCGACCA 3452
QY 3184 TTGTCATGCTCTGGGCTGTCCTCTGCTCTCTATCTTCTTCCCTTGTGGTTCGTTGTG 3243
Db 3453 TCGTCTGCTCTGGGCAATCTCTGCTGGGCTCATCTTCTCTTGTGTGGTTCGATCG 3512
QY 3244 ATCCATTCACCTACCGTCTCTGCTGGGCCCAAAATATCAAACTGTGGCATCACTGCTAGG 3303
Db 3513 ACCCTTCACCAACCGGCTCACTGCGCCGGATACCAGACGTTGTGGCATCACTGCTAGG 3572
QY 3304 AAAGTGGGAGTTT 3316
Db 3573 GAAGTGAAGGTT 3585

RESULT 10
US-10-160-719-5
; Sequence 5, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)...(3568)
US-10-160-719-5

Query Match
Best Local Similarity 70.9%; Score 2572.6; DB 15; Length 3773;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

QY 4 CGAGGAACCCCGCTCCAGCTCTGTGTCGGTTCGGGTTCGATCGCTCTGCCGCGCAATGG 63
Db 282 CCAGGCCCCAGGCTCCAGGCCAGCTCCCTCGACGCTTCTCGCGAGCTCGCTTCGCAATGG 341
QY 64 ACGGCGAGCGGAGCCCTGAGTCCGGAGGACACGGGGCCGGGACGTTGTGCCAGATCT 123
Db 342 AGGCGGACGCGGAGCGGCGTGAAGTCGGGAGGCGCGGTGGCGGACAGGTGTGCCAGATCT 401
QY 124 GCGCCGACGCGCTGGGGACCAACGCTTGGACGCGGACGCTTTCACCGCTGCGACGCTGCC 183
Db 402 GCGGCGAGCGGCTGGGGACCAACGCGGAGGGGACGCTTTCGCGGCTGCGACGCTGCGG 461
QY 184 GTTTCGCGTCTGCGGCCCTCTGTCTACGACACGAGCGCAAGAGGGGACCCAGGCTGCC 243
Db 462 GTTTCGCGTCTGCGGCCCTCTGTCTACGAGTACGAGCGCAAGGACGCGACGCGCTGCC 521
QY 244 TCCAGTCAAGACCAAGTATCAAGGCCACAGAGGAGGCCAGCATCCGCGGGAGGGAAG 303
```



```
QY 2464 CAGCTTTCAAGGATCTGCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTGCGGT 2523
Db |||
QY 2733 CAGCTTTCAAGGATCTGCCCCATCAATCTTTTCGACCGTCTGAACCAAGTCTGCGGT 2792
Db |||
QY 2524 GGGCTCTGGTCTCTGTTGAAATCTTTTCAGCGGCAITGCCCTTTATGTTATGTTAGTACG 2583
Db |||
QY 2793 GGGCTCTGGTCTGGTGGAGATCTCTTCAGCGGCACTGCCCTGTGTGTTAGTACG 2852
QY 2584 GAGGGCGCTCAAGTCTCTGGAGAGATTCGCTTACATCAACACCACTTTACCACTAA 2643
Db |||
QY 2853 GAGGGCGCTCAAGTCTCTGGAGAGATTCGCTTACATCAACACCACTTTACCACTAA 2912
QY 2644 CTTCTCTCCGCTCTAGTCTTATTTGATATTCGCTCTATCTGCTGCTCACTGGAAGT 2703
Db |||
QY 2913 CGTCCATCCGCTTCTCATCTACTGATCTCTGCGCCGCACTGCTGCTCAACCGGAAGT 2972
QY 2704 TCATCATCCAGAGATTTAGCAACTTTGGCCAGTATCTGTTGTTCAATGCGCTCTCTTTCAA 2763
Db |||
QY 2973 TCATCATCCAGAGATTCAGCAACTTTGGCCAGTATCTGTTGTTCAATGCGCTCTCTTTCAA 3032
QY 2764 TTTTCGCCACTGTTATCTTGGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2823
Db |||
QY 3033 TCTTCGCCACGGGCACTCTGGAGATGAGTGGAGCGGGTGGGCACTCGACGAGTGGTGA 3092
QY 2824 GGAATGAACAGTCTCTGGGTCATTTGAGGTATCTCTGCAATCTGTTGCGGCTTTTCAGG 2883
Db |||
QY 3093 GGAACGAGAGTCTTGGGTGATCGGGGCACTCTCGGCACTCTTCGCGGTGTTCCAGG 3152
QY 2884 GTCTTCTGAAGTGTCTCGCGTATCGACACCAACTTCACTGCTCACTCAAAAGGCTAATG 2943
Db |||
QY 3153 GCCTGCTCAAGTGTCTGGGCACTCGACACCAACTTCACTGCTCACTCAAAAGGCTAATG 3212
QY 2944 ACGAGAGAGGCACTTTCTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCCCTC 3003
Db |||
QY 3213 ACGAGAGAGGCACTTTCTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCCCTC 3272
QY 3004 CGACGACCAATTTGATCAATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3063
Db |||
QY 3273 CCACCACTCTGATCATCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3332
QY 3064 ACAGTGGTTACCAATCAATGAGGCGCTCTTTGGGAAGCTCTTTGTTGTTGTTGTTGTTG 3123
Db |||
QY 3333 ACAGGAGTACAGTCTGAGGCGGCTCTTCGGCAAGCTCTTCTTCGCTTCTGGTCA 3392
QY 3124 TTGTTCACTTATACCACTCTCAAGGCTCTTATGGGAGGCAAAACCGGACACCGGCA 3183
Db |||
QY 3393 TCGTCCACTCTACCCGTTCTCAAGGCGCTCATGGGAGGCGAGAACCGGACCCCGACCA 3452
QY 3184 TTGTCATCTCTGGGCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3243
Db |||
QY 3453 TCGTCTGCTCTGGGCACTCTGTTGGGCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3512
QY 3244 ATCCATTCATACCCGCTCTGCTGCGCCCAATATCCAACTGTGGCACTCACTGCTAGG 3303
Db |||
QY 3513 ACCCTTTCAACACCGGCTCACTGCGCGGATACCCAGCTGTGGCATCACTGCTAGG 3572
QY 3304 AAGTGGGAGTTT 3316
Db |||
QY 3573 GAAGTGGGAAGTTT 3585
```

RESULT 11

```
US-10-627-132-13
; Sequence 13, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Theoreof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
```

```
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-13
```

```
Query Match 64.2%; Score 2326.2; DB 13; Length 3704;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 2687; Conservative 0; Mismatches 543; Indels 9; Gaps 3;

QY 66 GCGCAGCGGAGCGCCCTGAAAGTCCGGGAGGACACGGGCGCGGAGACGTGTGCCAGATCTGC 125
Db |||
QY 275 GACGCGCGGAGCGCCACCAATTCGGGGAAGCATGTGGCGGCGGACGTGTGCCAGATCTGC 334
Db |||
QY 126 GCGCAGCGGCTGGGACCAACAGTTTGAACGCGGACGCTTTCACGCTGCGAGCTGTGCCGC 185
Db |||
QY 335 GCGCAGCGGCTGGGACCGCGGACGCGGACCTCTTTCACGCTGCGAGCTGTGCCGC 394
QY 186 TTCCCGGCTCGCGCCCTGTTACGACGACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 245
Db |||
QY 395 TTCCCGGCTGGCGCCCACTGTTACGAGTACGAGCGCAAGGAGGAGGAGGAGGAGGAGGAG 454
QY 246 CAGTCAAGACCAAGTACAAGCGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 305
Db |||
QY 455 CAGTCAAGACCAAGTACAAGCGCCACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 514
QY 306 GACGACCTGATGCGGATGATGAGTGTGATGCTTCACTACCTGTGATCTGTGGCACTGAGGAC 365
Db |||
QY 515 GAGGATGTGGATGTGTCAGATGTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 574
QY 366 CAGAAGCAGAAAGATTTGTCAGAGGATGCGGAGTGGGCGATGACATGACACCGGCGGAGTGC 425
Db |||
QY 575 CAGAAGCAGAAAGATTTGTCAGAGGATGCTCACTTGGCGGAGGAGGAGGAGGAGGAGGAG 632
QY 426 AATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAATGATGACAGTGA 485
Db |||
QY 633 -ATATTGGCTGGCTAAGTATGACAGCGGTGAATTTGGGATGGGAGTATGACAGTGGT 691
QY 486 GAGATCCCTAGGGATAGTCCCTTTCAGTACCAACAGCGGAGATGTTCAGGAGAAATCCCT 545
Db |||
QY 692 GAGATCCCTTCTGGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCT 751
QY 546 GGAGTCTCGCTGATCATCATCATGATGTCCTTACGCGGAGATCATCAGCAGAGTGTCTCCG 605
Db |||
QY 752 GGAGCTTCCCTGATCATCATCATGATGTCCTTGGGAACTTGGGAGGAGTGTGGGAGGAG 811
QY 606 TTTCCCTATGTGAATCATTCACCAATCCCTCAAGGAGGAGTGTCTCCGCGAGTATTTGGGAAT 665
Db |||
QY 812 TTTCCCTATGTGAATCATTCCTCAACCCATCGAGGAGGAGTGTCTCCGCGAGTATTTGGCAAT 871
QY 666 GTTCCCTGGAAAGAGAGAGTGTGCTGGAAATGAAGCAGGAGCAAGGAGTGTGCAATCCC 725
Db |||
QY 872 GTTGCATGGAAAGAGAGGAGTGTGATGGATGGAAATGA---AGGATAAAGGTGCAATCCCT 928
QY 726 ATGACTAATGGCAACAGTGTCTCCCTCTGAGGAGTGGGCGAGCTATGACATGATGCA 785
Db |||
QY 929 ATGACCAATGGCAACAGTGTCTCCATCAGAGGCGGTGGAGTGTGATATTTGATGCT 988
QY 786 TCTACTGAATACACATGGAAGAGCGCTTTTACTGAATGATGAAATCGCCAGGCTCTATCT 845
Db |||
QY 989 TCTACTGATTATACATGGAAGATGCTTTACTGAATGATGAAATCGCCAGGAGCTCTATCT 1048
```

QY 846 AGAAAGTCCCATTTCTTCCTCCAAATAAATCCCTACAGAAATGGTCAATTTCTGTGGG 905
D5 1049 AGAAAAGTGCCAAATTCCTTCATCCAGAAATAAATCCGTACAGAAATGGTCAATTTCTGTACGT 1108
QY 906 TTGGTTGTTCTAAGCATCTTCCTGCACTACCGTCTCCAAATCCCTGTGTGCGTAATGCATAC 965
D5 1109 TTGGCTGTTCTATGCATATCTTGTGGCTACCGGTATCACACATCTCTGTGAACAATGCATAT 1168
QY 966 CCATGTGTGCTTTATCTCTTATATGTAGATTTGGTTTGTCTTTATCTCTGTGATCTGGAT 1025
D5 1169 CCATGTGTGCTTTTATCTCTTATATGTAGATTTGGTTTGTCTTTGTCTGTGATTTTGGAT 1228
QY 1026 CAGTTCCTCCAGTGGTTTCCAAATCAACCGGAGACCTACCTGTATAGATGGCTTTAAGG 1085
D5 1229 CAGTTCCTCCAGTGGTTTCCAAATCAACCGGAGACCTACCTGTATAGATGGCTTTAAGG 1288
QY 1086 TATGACCGAGAGAGTGAACCGTCTCAGTTGGCTGTCTGTGACATATTTGTCTAGTACAGTC 1145
D5 1289 TATGACCGAGAGAGTGAACCATCTCAATAGTCTCTGTGATATTTTGTCTAGTACAGTC 1348
QY 1146 GACCCCTTGAAGAGAGCACTATCTGTCACCTGCAACACTGTGCTATCCATCTTGTGCTGT 1205
D5 1349 GATCCAAATGAAGAGAGCTCTCTGTCTGCTGCAAAATCTGTGCTTTCCATCTCTGTCTG 1408
QY 1206 GATTATCCCGTGACAAAGGTCTCTGTCTATGTATCTGTGACGAGCTTCAATGCTGACT 1265
D5 1409 GATTATCCCGTTGACAAAGGTATCTGTCTATGTATCTGTGACGAGCTGCTATGCTGACT 1468
QY 1266 TTTGAGCGCATGCTCTGAGACTTCCAGAGTTTGTCTAGGAAATGGGTACATTTGTGAAGAAG 1325
D5 1469 TTTGATGCTCTCTCTGAACCTTCAGAGTTTGTCTAGGAAATGGGTTCCTGTCTCTGAAGAAG 1528
QY 1326 TATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTGTCCAGAAATTTGATTAACCTGAAA 1385
D5 1529 TACAACATAGAGCTTAGGGCCCGGAATGCTACTTTGCTCAGAAATTTGATTAACCTGAAA 1588
QY 1386 GACAAAGTCCAGGCTTCATTTGTTTAAAGACCGCGGCCATGAAGAGAGAAATATGAAGAA 1445
D5 1589 GACAAAGTTCAAACCTCATTTTGTGAAGAACGCGGGCCATGAAGAGAGAAATATGAAGAA 1648
QY 1446 TTTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCAATGAAGTCCCGAGGAGAGATGG 1505
D5 1649 TTTCAAAGTTCGTATCAATGGTCTGTAGCCAAAGGCAAAAAAGTTCCCGAGGAGGATGG 1708
QY 1506 ATCATGCAAGATGGCACACCATGCCAGGAGAAACATACACGAGGATCATCTCTGGAATGATT 1565
D5 1709 ATCATGCAAGATGGTACACCTTGGCCCTGGGAAACATACAGGACCATCTCTGGAATGATT 1768
QY 1566 CAGTTTTCCTTGTGTACAGTGTGGCTTGTATCTAGGAGTAAATGAGTCCCGGTTTA 1625
D5 1769 CAGTTTTCCTTGTGTGTACAGTGTGGCTTGTATCTAGGAGTAAATGAGTCCCGGTTTG 1828
QY 1626 GTTATATGTCTCGTGAAGAGCTCTGGGTTCCAGCACCAAGAAAGGCTGTGGCCATG 1685
D5 1829 GTTATATGTCTCGTGAAGAAAGCTCTGGATTCCAACATCACAGAAAGGCTGTGGCCATG 1888
QY 1686 AATGCCCTTGTCTGTGTCTAGCTCTCTTACTAATCGACAAATACATGTTGAATCTTTGAT 1745
D5 1889 AATGCACTTGTGTGTGTGTCTCTTACTAATGGGCAATACATGTTGAATCTTTGAT 1948
QY 1746 TGTGATCACTACATCAACACAGCAAGGCTGTCCGAGAGCTATGTGCTTCTCTAATGGAT 1805
D5 1949 TGTGACCACTACATCAATATAGCAAGGCTTCCGAGAGCTATGTGCTTCTCTAATGGAC 2008
QY 1806 CCAAACTAGTCCGCAAGTCTGTATGTGAGTTCCTCCAGAAAGGCTGTGGATGAT 1865
D5 2009 CCAAACTAGGAGGATGTCTGTATGTCCAAATTTCTCCAGAGGTTTGTATGATGAT 2068
QY 1866 AGGATGATCGATATGCAACAGGACACTCTCTTTTGTATATTAACCTCAGGCGCT 1925
D5 2069 AGGAATGACCGATATGCAACAGGACACTCTGTGTCTTTCGATATTAACCTCAGGAGCT 2128
QY 1926 GACGGCATTTCAAGGACCAGTTTATGTGGGAACTGGTTGTGTTTTCAACAGAACAGCTATC 1985

D5 2129 GACGGCATTTCAAGGCGCAGTTTATGTGGGAACGTGTGTGTGTATTAACAGAACGGCTTA 2188
QY 1986 TATGTTTATGAGCCCCCAATTAAGCGAAGAGCCAGGTTTCTTGTGCATCACTATGTGGG 2045
D5 2189 TATGTTTATGAGCTTCCAGTCAAGAAAAAAGAGCCAGGCTTCTTCTTGTGCTTGTGGG 2248
QY 2046 GGCAGAGAAAGCGCAAGCAAGTCAAGAAAAAGAGCTCAGATAAGAAAAAGTCGAACAAG 2105
D5 2249 GGAAGGMAAAGACGTCAAAAATCTAAGAA--GAGCTCGAAAAAGAAAGTCAACATAGA 2305
QY 2106 CATGTGGAACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGCTGTGAGGT 2165
D5 2306 CACGACAGACAGTTCTGTACCACTATTATCTCGAAGATATAGAGAAAGGATTTGAAGGT 2365
QY 2166 GCTGGTGTGTATGATGAGAAAAATCAGTTCTCATGTCTCAATGAGCTTAGAAGAGATTT 2225
D5 2366 TCTCAGTTTGTATGATGAGAAATCGCTGATTATGTCTCAATGAGCTTGGAGAAGATTT 2425
QY 2226 GGCAGTCAAGCAATTTGTTGCTCCACTCTGATGAATAATGGTGTGTTCTCAGTCC 2285
D5 2426 GGCAGTCCAGTGTGTTTGTAGCTCTACTCTGATGGAATATGGTGTGTTCCACAATCT 2485
QY 2286 TCCACTCCAGATCTCTTTTGAAGAGCTATCCATGTCTAAGTTGTGCTATGAGGAC 2345
D5 2486 GCAACTCCAGAGTCTCTCTGAAAGAGCTATTTCATGTCTCATGCTGTGGCTATGAGGAC 2545
QY 2346 AAGTCTGAAATGGGAATCAGATTGGTGTGATCTATGGATCTCTCACAGAAGATATTCTT 2405
D5 2546 AAACTGACTGGGGAATCAGATTGGTGTGATCTATGGTCTTACAGAGACATCTC 2605
QY 2406 ACTGATTTCAAGATGCAACGAGAGCTGGCGTTCTATGTCATTTGTCATGTCGATGCGGCCA 2465
D5 2606 ACCGATTTCAAGATGCAATGCTCGAGGCTGGCGATCAATCTACTGCTATGCTAAGCGACCA 2665
QY 2466 GCTTTCAAGGATCTGCCCCCATCAATCTTTCAGATGCTCTGACCAAGGCTGCGGTGG 2525
D5 2666 GCTTTCAAGGATCTGCTCTCTATCAACCTTTCGATCGTTTGAATCAAGTGTCTCGGTGG 2725
QY 2526 GCTCTCGTCTCTGTGAATTTCTTTTTCAGCGGCTATGCCCCCTTATGTTGTTGCTACGGA 2585
D5 2726 GCTCTTGTGTTCCATGAAATCTTTTTCAGCAGGCTGTGACCAAGGCTGCGGTGG 2785
QY 2586 GGGCGCTCAAGTTCCTGGAGAGATTCGTTTACATCAACACCACTATTAACCACTAAC 2645
D5 2786 GGGCGCTTAAATTTCTGGAGAGATTTGCTTATCAACACCACTATTAACCACTAAC 2845
QY 2646 TCTCTCGGCTTCTAGTCTATTGTATTTGCGCTCTATCTGTCTGCTCACTGGAAAGTTC 2705
D5 2846 TCAATCCGCTCTCTGTACTGCATATGCCAGCAGTTTGTCTCTCACTGGGAAGTTC 2905
QY 2706 ATCATGCCAGATTTAGCAACTTTGGCCAGTATCTGTTTCATTTGGCTCTTCTTCAATT 2765
D5 2906 ATCATCCCAAGATTTAGTAACCTAGAGGTTTGGTTTATATCGCTCTTATCTCAATC 2965
QY 2766 TTGCCACTCGTATCTCTTGAAGATGAGTGTGAGTGTGTTGGCATGACGAGTGTGAGG 2825
D5 2966 TTTGCCACTCGTATCTCTTGAAGATGAGTGTGAGTGTGTTGGCATGACGAGTGTGAGG 3025
QY 2826 AATGAACAGTTTGGGTCATTTGAGGTTATCTCTGCACATCTGTTTGGCTCTTTCAGGT 2885
D5 3026 AACGAGCAGTTTGGGTCATTTGAGGTTATTTCTGCGCATTTATTTGCGCTTCTCCAGGT 3085
QY 2886 CTTCTGAAGTGTCTGCGGTATCGACCACTTCTCACTGTCACTCAAAAGGCTTAATGAC 2945
D5 3086 CTCCTGAAGTGTCTGCTGTGTATCGACACGAGCTTCACTGTCACTCTAAGGCCACTGAC 3145
QY 2946 GAAGAAGCGCATTTGTCTGAGCTCTACATGTTTCAAGTGGAGAGAGCTTCTCATCTCCCTCG 3005
D5 3146 GAAGAAGTGTATTTTGGCGGACTCTACATGTTTCAAGTGGAGAGAGCTTCTCATCTCCCTCG 3205
QY 3006 ACGACCATTTGATCATTTAATGATGTTGGTGTCTGCTGGCACTCTCTAGCCCATCAAC 3065

Db 3206 ACCACTATTTTGATCATCAACCTGCTCGGTGCTGGCATTTCTTCAACGAATCAAT 3265
Qy 3066 AGTGGTTACCAATCATGGGGCGGCTTTTGGAGCTCTTTCTTGCCCTTCTGGGTGATT 3125
Db 3266 AGCGGTTACCACTCATGGGACCTCTTTTCGGGAAGCTCTTCTTGCGTTCCTGGGTGATT 3325
Qy 3126 GTTCACTTATACCCATTTCTCAAGGCTCTTATGGGAGGCAAAACCGCACACCGAGATT 3185
Db 3326 GTCCACCTGTACCCCTTCTCAAGGCTCTATGGGAGACAGAACCGACGCGACCAAT 3385
Qy 3186 GTCATGCTGGGCTGCTCTCTCGCTTCTATCTTCTTCTGCTGCTGGGTTCTGTTGAT 3245
Db 3386 GTCTGTGCTGGCTATCTCTCTGCTGCTATCTTTCCTGATGTGGGTTCTGTTACGAT 3445
Qy 3246 CCATTCACCTACCGCTCTCGCTGGCCCAATATCCAACTGTGGCATCACTGCTAGGA 3304
Db 3446 CCATTCACACCCGGGTCACTGGCCCTGATATCGGAAATGTGGCATCACTGCTAGGA 3504

RESULT 12

US-10-209-059-13

; Sequence 13, Application US/10209059

; Publication No. US2003016383A1

; GENERAL INFORMATION:

; APPLICANT: Dhugga, Kanwarpal S.

; APPLICANT: Wang, Haiyin

; TITLE OF INVENTION: Maize Cellulose Synthases and Uses

; TITLE OF INVENTION: Theeof

; FILE REFERENCE: 0864R2

; CURRENT APPLICATION NUMBER: US/10/209,059

; CURRENT FILING DATE: 2002-07-31

; PRIOR APPLICATION NUMBER: 60/096,822

; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: 09/371,383

; PRIOR FILING DATE: 1999-08-06

; PRIOR APPLICATION NUMBER: 09/550,483

; PRIOR FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 3704

; TYPE: DNA

; ORGANISM: Zea mays

US-10-209-059-13

Query Match 64.2%; Score 2326.2; DB 15; Length 3704;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 2687; Conservative 0; Mismatches 543; Indels 9; Gaps 3;

Qy 66 GCGACGCGGACGCGCTGAAGTCCGGGAGGACGCGGCGGGGAGCTGTGCCAGATCTGC 125
Db 275 GACGGCGGACGCCAGAAATTCGGGAGAGCATGTGGCGGCGGAGGTGTGCCAGATCTGC 334
Qy 126 GCCGACGCGCTGGGACACACGTTGGACGCGACGCTCTTACCGCTCGGAGCTGTCCCGC 185
Db 335 GCGACGCGGCTGGGACACGCGGCGGACGCGACCTTTACCGCTCGGAGCTGTCCGCG 394
Qy 186 TTCCGGCTCTCCGCGCTGTACGACGAGCGAGGAGGACCCAGCGCTGCGCTC 245
Db 395 TTCCCGGTGTCCGCGCTGTACGAGTACGAGGCGAGGACGCGACCCAGCGCTGCGCG 454
Qy 246 CAGTGCAGACCAAGTACAAGCGCACAGAGGAGCCAGCGATCCGCGGAGGAGGC 305
Db 455 CAGTGCAGACTAGTACAGCGCCACAAAGGAGCCACAGTACACGGTGAAGAAAT 514
Qy 306 GACGACACTGTATGCCGATGATGGTAGTGACTTCAACTACCTGCACTTGGCACTGAGAC 365
Db 515 GAGGATGGGATGTGACGATGTGAGTGACTTCAACTACCAAGCATCTGGCAACAGGAT 574
Qy 366 CAGAGCAGAGATGTGTGACGAGATGCGGAGCTGGGCGCATGAAACAACGGGGGAGTGGC 425
Db 575 CAGAAGCAAAAGATTGTCTGAGAGATGTCTACTTGGCGGCAAAACTCACGTGGCAGTG- 632

Db	1709	ATCATGCAAGATGGTACACCTCGCCCTGGGAACAATACTAGGCACCATTCTGGAATGATT	1768
Qy	1566	CAGGTTCCTTTCCGTGACAGTGCTGGGCCCTTGATACTAGCGGTAATGAGCTCCCCCCTTTA	1625
Db	1769	CAGGTTCCTGGGTCACTGGAGGGCTTGAAGTTGAAGGCAATGAATTCCTCCTTTG	1828
Qy	1626	GTTTTATGTGTCGTGTAAAGCGTCTCGGTCTCCAGCACACAAGAAGGTGGTGCATG	1685
Db	1829	GTTTTATGTGTCGTGTAAAAACGTCTCGGATCCACATCAACAAGGCTGGTGCATG	1888
Qy	1686	AATGCCCTTGTTCTGTGTCCTCAGCTGTCTTAATAATGGAACAATACATGTTGAATCTTGAT	1745
Db	1889	AATGCACTTGTTCTGTGATCAGCTGTCTTTACTAATGGGCAATACATGTTGAATCTTGAT	1948
Qy	1746	TGTGATCACTACATCAACAAGAGGCTGTCGAGAACTATGTGCTTCCTTAATGGAT	1805
Db	1949	TGTGACCACTACATCAATAATAGCAAGGCTCTTCGAAAGCTATGTGCTTCCTTATGGAC	2008
Qy	1806	CCAAACCTTAGTCCGCAAGCTGTGTTATGTGCAAGTCCCAAAAAGGTTTGATGGGATGAT	1865
Db	2009	CCAAACCTAGGAAGGATGTCGTATGTCCCAATTTCTCAGAGGTTTGATGGTATGAT	2068
Qy	1866	AGGAATGATCGATATGCAAAACAGAAACAATGTCCTTTTTGATATTAACCTTGAGGGCCCTT	1925
Db	2069	AGGAATGACCGATATGCAAAACAGAAACAATGTCCTTTTTGATATTAACCTTGAGGGCTCTT	2128
Qy	1926	GACGGCAATCAAGAACACGTTTATGTGGGAACGCTGTTGTTTTCACAGAAACAGCTATC	1985
Db	2129	GACGGCAATCAAGGCCAGTTTATGTGGGAACGCTGTTGTTTTCACAGAAACGCGCTTA	2188
Qy	1986	TATGTTATGAGCCCCAAATTAAGCGGAAGACAGGCTTCTTTGGCATCACTATGTGGG	2045
Db	2189	TATGTTATGAGCCTCCAGTCAAGAAAAAAGACAGGCTTCTTCTTCGCTTTGTGGG	2248
Qy	2046	GGCAAGAAAGGCAAGCAAGTCAAAGAAAAAGAGCTCAGATAGAAAAAGTCCAACAAG	2105
Db	2249	GGAGGANAAGACGTCANAAATCAAGAA--GAGCTCGGAAAAAGAAAGTCACATAGA	2305
Qy	2106	CATGTGGACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTTGTGAAGT	2165
Db	2306	CACGCAGACAGTTCTGTACCAAGTATTTAATCTCGAAGATATAGAGGAAGGATGTAAGT	2365
Qy	2166	GCTGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAATATGACCTTTAGAGAAAGATTT	2225
Db	2366	TCTCAGTTTGATGATGAGAAATCGCTGATTAATGCTCAATGAGCTTTGGAGAGAGATTT	2425
Qy	2226	GGCCAGTCAGCAGCATTTGTTGCGCTCCACTCTGATGGAATATGGTGGTTCTCAGTCC	2285
Db	2426	GGCCAGTCAGTGTGTTGTAGCCCTCTACTCTGATGGAATATGGTGGTTCCCAATCT	2485
Qy	2286	TCCATCCAGAAATCTCTTTTGAAGAAGCTATCATGTCAATAGTTGTGGCTATGAGGAC	2345
Db	2486	GCAATCCAGAGTCTCTTCTGAAAGAGCTATTCATGTCATCAGCTGTGGCTATGAGGAC	2545
Qy	2346	AACTCTGAATGGGAAC TGAGTTGGTGGATCTATGGATCTGTGTCAGAGAATATTTCTT	2405
Db	2546	AAAAC TGA CTGGGAAC TGAGTTGGTGGATCTATGGTCTGTTACAGAAGACATTTCTC	2605
Qy	2406	ACTGGAATCAAGATCAAGCAGAGGCTGGGCTTCAGTCTATTGCATGCCCAAGCGCCA	2465
Db	2606	ACCGAATCAAGATCATGCTCGAGGCTGGCGATCAATCTACTGCATGCTCAAGCGACCA	2665
Qy	2466	GCTTTCAAGGGNATCGCCCCCATCAATCTTTTCAGATCGCTGAAACAAAGTCTCGGTGG	2525
Db	2666	GCTTTCAAGGGATCTGCTCTCAATCAACCTTTTCGGATCGTTTGAATCTCAAGTCTCGGTGG	2725
Qy	2526	GCTCTCGGTTCTGTTGAAATCTTTTCAGCGGCAATTCGCCCTTATGTAATGGCTACGGA	2585
Db	2726	GCTCTTGGTTCCATTTGAAATCTTTTCAGAGGCAATGTCCCATATGTAATGGCTATGGA	2785
Qy	2586	GGCGGCCCTCAAGTTCCTTGAGAGATTCGTTTACATCAACACCAATTTACCCATAACC	2645

```

RESULT 13
US-10-160-719-57
; Sequence 57, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS

```


QY	666	GTTCCTGGAAGAGAGAGTTGATGGCTGGAATGAAGCAGGACAAGGGTGGATTCCC	725
Db	732	GTTCATGGAAGAGAGGGTGGATGGAAATGA---AGATAAAGTGAATTCCT	788
QY	726	ATGACTAATGGGCAAGAGATTGCTCCCTCTGAAGGTCGGGCAGCTACTGACATCGATGCA	785
Db	789	ATGACCAATGGAAACAAGCAATGCTCCATCAGAAGGGCGTGGAGTTGCTGATATTGATGCT	848
QY	786	TCTACTGAAATCAACATGAAGACGCTTTACTGAAATGATGAACACTCGCAGCCTCTATCT	845
Db	849	TCTACTGATTAATGACATGAAGATGCTTACTGAATGATGAACACTCGGCAACCTCTATCT	908
QY	846	AGAAAAGTCCCATTTGCTTCCCTCCAAAATAAATCCCTACAGAAATGCTATTGTTCTGCGG	905
Db	909	AGAAAAGTCCCAATTCCTTATCCAGATTAATCCGTACAGAAATGCTATTGTTGCTACGT	968
QY	906	TTGGTTGTTCTAAGCATCTTCCTGCACTACCGTCTCACAAATCCCTGTGCGTAAATGCATAC	965
Db	969	TTGGCTGTTCTATGCATATTCCTTGCGCTACCGTATCCAGTATCCATGGAACATGCAAT	1028
QY	966	CCACTGTGGCTTTATCTGTTATATGTGAGATTTGGTTTGGTTTATCTCTGGATACTGGAT	1025
Db	1029	CCACTGTGGCTTTTATPCCGTCAATATGTGAGATCTGGTTTGGTTTGTCTCTGGATTTGGAT	1088
QY	1026	CAGTTCCCGAAGTGGTTTCCAAATCAACCGGGAGACCTACTCTTGATAGACTGGCTTTAAGG	1085
Db	1089	CAGTTCCCAAGTGGTCCCCATCAACCGTGAACATACCTTGATAGACTGGCTTTAAGG	1148
QY	1086	TATGACCGAAGAGGTGAACCGTCTCAGTTGGCTGCTGTGATGATATTTGTCACTAGTACAGTC	1145
Db	1149	TATGACCGAAGAGGTGAACCATCTCAATTAGCTCTGTTGATATTTTGTCACTAGTACTGTG	1208
QY	1146	GACCCCTTGAGAGGACCACTATCGTCACTGCCAACACTGTGCTATGCATTCCTGCTGTT	1205
Db	1209	GATCCAAATGAAGGACCTCTCTCTGCTCACTGCAAAATCTGTGCTTTCCATCCTTCTGCTC	1268
QY	1206	GATTATCCCGTGGCAAGGTCTCTTGTCTATGATCTGATGAGGAGCTTCAATGCTGACT	1265
Db	1269	GATTATCCCGTGTACAAAGTATCTTGCTATGTTTCGGATGATGAGAGCTGTATGCTGACT	1328
QY	1266	TTTGACGATTTGGGTGAGACTTACAGATTGCTAGGAAATGGGTACCAATTTGTGAAGAAG	1325
Db	1329	TTTGATGCTCTCTCTGAAATCTACAGATTGCTAGAAATGGGTTCCGTTCTGTGAAGAAG	1388
QY	1326	TATGACATTTGAACCAAGCTCCGAGTTTACITTTTGCAGAAATTTGATTAACCTGAAA	1385
Db	1389	TACAAATAGAGCCTTAGGGCCCCGAATGGTACTTTGCTCAGAAATTTGATTACTTGAAA	1448
QY	1386	GACAAAGTCCAGCCTTCATTTTGTAAAGACCGCGGCCCATGAAGAGAGAATATGAAGA	1445
Db	1449	GACAAAGTTCAAACTCATTTGTGNAAGAACCCCGGCCATGAAGAGAGAATATGAAGA	1508
QY	1446	TTTAAATACAGATAAATGCCCTAGTTTCTAAGGCATTTGAAGTCCCGCAGGAAAGATGG	1505
Db	1509	TTCAAAGTTCGTATCAATGGTCTTTAGCCCAAGGCACAAAAGTTCCCGCAGGAGGATGG	1568
QY	1506	ATCATGCAAGATGCAACACATGCCCCAGCAACATACAGGGATCATCTCTGGAATGATT	1565
Db	1569	ATCATGCAAGATGATACACCTTTGGCCCTGGGAACAATACTAGGGACCATCTCTGGAATGATT	1628
QY	1566	CAGGTTTTCCTTGCTCAGTGGTGGCCTTGATATCTGAGGGTAATGAGCTCCCGGTTTA	1625
Db	1629	CAGGTTTTCCTTGCTCAGTGGGAGGCTTGACGTTGAGGCAATGAATCTTCTCGTTTG	1688
QY	1626	GTTTATGTGCTCGTGAAGACGCTCTGGTTTCCAGCACCAAGAGGGCTGGTGCCATG	1685
Db	1689	GTTTATGTGCTCGTGAAGAACGCTCTGATTTCCAAATCACAGAGGGCTGGTGCCATG	1748
QY	1686	AATGCCCTTGTGCTGCTGAGCTGCTCTTACTAATGAGCAATACATGTTGAATCTTGAT	1745
Db	1749	AATGCACTTGTGCTGATGAGCTGCTCTTACTAATGGGCAATACATGTTGAATCTTGAT	1808

QY	1746	TGTGATCACTATCATCAACACACAGCAGGCTGTCCGAGAAGCTATGTGCTTCTTAATGGAT	1805
Db	1809	TGTGACCACTATCATCAATTAATAGCAAGGCTCTTCGAGAAGCTATGTGCTTCTTATGGAC	1868
QY	1806	CCAAACCTAGGTCCGCAAGTCTGTTATGTGCAAGTTCCCAACAAAGTTTCATGCGGATTGAT	1865
Db	1869	CCAAACCTAGGAAGGATGCTGTTATGTGCAATTTCTCAGAGGTTTGATGTTATGAT	1928
QY	1866	AGGAATGATTCGATATGCAAAACAGGAACACTGTCTTTTGTATTAATTAACCTTGAGGGCCCTT	1925
Db	1929	AGGAATGACCGATATGCAAAACAGGAACACTGTGTTTTCGATATTAACCTTGAGAGCTCTT	1988
QY	1926	GACGCAATTCAGGACCACTTTATGTGGGAACCTGTTGTTTCTCAACAGAACAGCTATC	1985
Db	1989	GACGCAATTCAGGACCACTTTATGTGGGAACCTGTTGTTTAAACAGAACCGGCTTTA	2048
QY	1986	TATGTTATGACGCCCCCAATTAAGCGGAAGAGCAGGTTCTTGGCATCATATGTGGG	2045
Db	2049	TATGTTATGACGCTCCAGTCAAGAAATAAGCAGGCTTCTTCTTCGCTTTGTGGG	2108
QY	2046	GGCAAGAAAGAGGCAAGCAAGTCAAGAAAGAGAGCTCAGATPAAGAAAAGTCGAACAAG	2105
Db	2109	GGAGGAAAAGAGCCTCAAAATCTAAGAA---GAGCTCGAAAAGAGAGTCAATAGA	2165
QY	2106	CATGTGGACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGGTGTGAAAGT	2165
Db	2166	CACGACAGCTTCTGTACCACTATTTAATCTCGAAGATATAGAGGAAGGATTTGAAAGT	2225
QY	2166	GCTGGTTTGTATGATGAGAAATCAGTTCTCATGCTCAAAATGAGCTTAGAGAGAGATTT	2225
Db	2226	TCTCAGTTTGTATGATGAGAAATCGCTGATATGCTCAAAATGAGCTTGGAGAAGAGATTT	2285
QY	2226	GGCAGTCAAGCAGATTTTGGCTCCACTCTGTATGGAATATGTTGTTCTCCTCAGTCC	2285
Db	2286	GGCAGTCCAGTGTGTTGTAGCCCTACTCTGATGGAATATGTTGTTTCCCAACT	2345
QY	2286	TCCACTCCAGATCTCTTTTGAAGAGCTATCATGTCATAGTTTGGCTATGAGGAC	2345
Db	2346	GCAATCCAGAGTCTCTTCTGAAAGAGCTATTCATGTCATCAGCTGTGGCTATGAGGAC	2405
QY	2346	AAGTCTGAATGGGAATCGAGATGTTGGATCTATGATCTGTCAACAGAGATATTTCTT	2405
Db	2406	AAACTGACTGGGAACTGAGATTTGGTGGATATATGTTCTGTTACAGAGACATTTCTC	2465
QY	2406	ACTGATTCAGAGTGCACGCAAGAGGCTGGGCTTCACTATTTGATGCCAACGCCCA	2465
Db	2466	ACCGAATCAAGATGCATGCTCGAGGCTGGCATCAATCTACTGATGCTTAAGGACCA	2525
QY	2466	GCTTTCAGGAGTCTGCCCCCATCAATCTTTCAGATCGTCTGAACCAAGTGCTGGGTGG	2525
Db	2526	GCTTTCAGGAGTCTGCTCTCTATCAACCTTTCCGATCGTTTGAATCAAGTGCTTGGGTGG	2585
QY	2526	GCTCTCGGTTCTGTTGAAATCTTTTTCAGCGGCAATTCGCCCTTATGTTATGGCTACCGA	2585
Db	2586	GCTCTGTTCCATTTGAAATCTTTTTCAGCAGGCAATTTGCCATATGTTATGGCTATGGA	2645
QY	2586	GGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTTTACCACTAAC	2645
Db	2646	GGCGGCTTAAATCTCTGGAGAGATTTGCTTATATCAACACCAATTTATCCACTCACA	2705
QY	2646	TCTCTCCGCTTCTAGTCTATTTGATATGCTGCTCTATCTGTCTGCTCACTGGAAAGTTC	2705
Db	2706	TCAATCCGCTCTCTCTGTTATCTGCATATTTGCCAGCTTTTGTCTTTTCACTGGGAAGTTC	2765
QY	2706	ATCATGCCAGATTTAGCAACTTGGCCAGATCTGTTTCAATGCGCTCTTCTTCTTCAAT	2765
Db	2766	ATCATCCCAAGATTTAGTAACTTAGAGAGTGTGTTGGTTTATATCTGCTCTTTATCTCAAT	2825
QY	2766	TTGCCACTGGTATCTTGGAGATGAGGTGGAGTGTGTTGGCATTGACGAGTGGTGAGG	2825
Db	2826	TTTGCCACTGGTATCTTGGAGATGAGGTGGAGTGTGTTGGCATTGATGATGTTGGAGG	2885
QY	2826	AATGAACAGTTTGGGTCTATGAGGATCTCTGCAATCTGTTTGGCTCTTTCAGGGT	2885


```
QY 1423 CCATGAAGAGAGATATGAAGAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCAT 1482
Db |||||
QY 2130 CCATGAAGAGGAAATAGAGAAATTAAGTTCTGGTTCTCGTCGCAAGGCAC 2071
Db |||||
QY 1483 TGAAGTCCCGAGGAAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATA 1542
Db |||||
QY 2070 AGAAAGTTCTGAGGAAGGATGATATGCAAGATGGCACACCTTGGCTGGTGAACAATA 2011
Db |||||
QY 1543 CCAGGATCATCTCGAATGATCAGGTTTCCCTTGGTCAAGTGGTGGCCCTTGATCTG 1602
Db |||||
QY 2010 CCAGAGACCATCTCGGAATGATTCAGGTTTCCCTGGTCTAGTGGAGGCCCTTGATACCG 1951
Db |||||
QY 1603 AGGGAATAGAGCTCCCGTTTGTATGTTCTCTCGTGAAGAGCGTCTCGGTTCCAGC 1662
Db |||||
QY 1950 AAGCAATAGAGCTTCTCGTTAGTCTAIGTCTCTGAGAAACGTCCTGATTTCCAA 1891
Db |||||
QY 1663 ACCAAGAAAGGCTGGTCCCATGAATGACCTTGTTTCGTCTCAGCTGTCCCTTACTAATG 1722
Db |||||
QY 1890 ACCATAAAAGGCTGGTCTATGAATGACCTTGTTTCGTGTATCAGCTGTCCCTTACTAATG 1831
Db |||||
QY 1723 GACAAATACATCTGAATCTTGAATGATCACTACATCAACAAAGAGGCTGTCCGAG 1782
Db |||||
QY 1830 GACAAATCTTGTGAATCTTGAATGATCACTACATCAACAAATAGCAAAAGCTCTCCGAG 1771
Db |||||
QY 1783 AAGCTATGTGCTTCTTAATGATCCAAACCTAGGTCCGCAAGTCTGTTATGTGCAAGTTCC 1842
Db |||||
QY 1770 AGGCTATGTGCTTCTTATGATCCAAACCTAGGAGCGGTCTGTTTATGTCCAAATCC 1711
Db |||||
QY 1843 CACAAAGTTTGAATGGGATGATAGGAATGATCGATATGCAAAACAGAAACATGTTCTTTT 1902
Db |||||
QY 1710 CTCAGAGTTTGAAGGTTATCGATAGAAATGATCGATATGCAAAACAGAAACACCGGTGTTT 1651
Db |||||
QY 1903 TTGATATTAATCTCAGGGGCTTGAAGGATTCAGGACATTCAGGACAGTTTATGTGGAACTGGTT 1962
Db |||||
QY 1650 TCGATATTAATTTGAGAGGTTCTCGATGGTCTCCAAGGACCAAGTTTATGTGGGAACCGGTT 1591
Db |||||
QY 1963 GTGTTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAGSCGAAGAGCCAG 2022
Db |||||
QY 1590 GTGTGTTCAACAGAACAGCTCTTATGTTTATGAAACCCCCCATTAAGCAGAAGAGGCCAG 1531
Db |||||
QY 2023 GTTTCCTTGGCATCATATGTGGGGCAAGAAAGAGGCAAGCAAGTCAAAAGAAAAGGAGCT 2082
Db |||||
QY 1530 GTTACTTCTCTCTGCTTGTGGGGACGAAAGAGAGCAAAAAGTCCGAAGGAGAGACA 1471
Db |||||
QY 2083 CAGATAGAAAAGTCCGAACAGCATGTGGACAGTTCTGTTTCCAGTATTAATCTCGAAG 2142
Db |||||
QY 1470 CGGAAAAGAAAAGTCAACAAAATGTGGACAGTTCTGTGCCAGTTTTTAATCTTGAAG 1411
Db |||||
QY 2143 ACATAGAGGAGGGTGTGAAGTGTCTGGGTTTGTATGATGAGAAATCAGTTCTCATGTCTC 2202
Db |||||
QY 1410 ATATAGAGGAAGGATGAAGTTCTGGAATTTGATGAGAAATCACTCATGATGCTC 1351
Db |||||
QY 2203 AATAGGCTTAGAAGAGATTTGGCCAGTCAGAGCATTTGTTGCCCTCCACTCTGATGG 2262
Db |||||
QY 1350 AATAGGCTTAGAAGAAAGATTTGGTCAAATCTAGTGTGTTTGTAGCCCTCCACTCTGATGG 1291
Db |||||
QY 2263 AATATGTTGTTTCTTCCAGTCCCTCCACTCAGAAATCTCTTTTGAAGAAAGCTATCCATG 2322
Db |||||
QY 1290 AATATGTTGTTTCTTCCAAATCCGCAATCCAGAAATCTCTTCTGAAAAGAGGCCATACATG 1231
Db |||||
QY 2323 TCATAAGTTGTGGCTATGAGGACAAGTCTGAATGGGAACATGAGATGGTTGGATCTATG 2382
Db |||||
QY 1230 TTATCAGCTGTGGCTATGAGACAAAAGTGAATCTGGGAATCTGAGATTTGATGATCTATG 1171
Db |||||
QY 2383 GATCTGTCAAGAAAGATATTCTTACTGGATTAAGATGCAAGGAGGCTGGGCTTCCAG 2442
Db |||||
QY 1170 GTTCTGTACAGAAGATATTCTCACTGGATTCAGATGCAATGCACTGGCTGGCTGGCGGTCAA 1111
Db |||||
QY 2443 TCTATTCATGCCCAAGGCCAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAATC 2502
Db |||||
QY 1110 TCTACTCATGCTTAAAGCGACCAAGCATTTCAAGGGGTCTGCCCCCAATTAATCTTTCCGATC 1051
Db |||||
```

```
QY 2503 GTCTGAACCAAGTCTGCTGGGTGGGCTCTCGGTTCTGTTGAATTTCTTTTCAGCCCGGCATT 2562
Db |||||
QY 1050 GTCTGAACCAAGTCTGCTGGGTGGGCTCTGTTCTGTTGAATTTCTTTTAGCAGGCATT 991
Db |||||
QY 2563 GCCCTTATGTGTATGGCTTACGAGGCGGCTCAAGGTTCTCTGGAGAGATTTCGCTTACATCA 2622
Db |||||
QY 990 GTCCCATATGTTATGGGTACGGAGACGGCTTAAGTTCTCTGGAGAGATTTCGTTACATCA 931
Db |||||
QY 2623 ACACCAATTTACCCCAATACCTCTCTCCGCTTCTAGTCTATTTGATATATTCGCTGCTA 2682
Db |||||
QY 930 ACACCACTATTTATCCACTCATCGATCCACTCTCTTGTACTGCATATATTCGCAGCTA 871
Db |||||
QY 2683 TCTGTCTGCTCACTCGGAAGTTTCATGACGACAGATTAGCAACTTTGGCCAGTATCTGTT 2742
Db |||||
QY 870 TCTGTCTTCTCACTGGGAAGTTTCATACACGAGATTAGCAACTTTTCGAGTATTTGTT 811
Db |||||
QY 2743 TCATTTGCGCTCTTCTCTTTCAATTTTCGCACTCTGGTATCTTTGAGATGAGGTGGAGTGGT 2802
Db |||||
QY 810 TTAATCTCTCTCTTCTGTCATCTTTCGCTACTGTTATCTTGAGATGAGGTGGAGTGGT 751
Db |||||
QY 2803 TTGGCATTTGACGAGTGTGGAGGAATGAACAGTTCTGGGTCAATTTGAGGTATCTCTGCA 2862
Db |||||
QY 750 TTGGCATCGATGAATGTGGAGGAATGAGCAGTTTGGGTTATTTGAGGTATTTCCGCTC 691
Db |||||
QY 2863 ATCTGTTTCCGCTCTTTTCAAGGTTCTTGAAGGTGCTTTCGCGGTATCGACACCACTCA 2922
Db |||||
QY 690 ATCTATTTCCGCTCTTCCAGGTTCTGCTGAAGGTGCTTGTGGAAATTTGATATCCAGCTTCA 631
Db |||||
QY 2923 CTGTCACTTCAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTCTACATGTTTCAAGT 2982
Db |||||
QY 630 CTGTCACTTCAAGGCTCTGATGAAGAGGCGATTTGCTGAGCTCTACATGTTTCAAGT 571
Db |||||
QY 2983 GGACGAGCTTCTCATCCCTCCGACGACCAATTTTGAATTAACATGTTGGTGTGTTG 3042
Db |||||
QY 570 GGACAAACCTTCTGATCCCAACCGACCACTATTTTGAATCATCAACCTGTGTGGTGTGTTG 511
Db |||||
QY 3043 CTGGCACTCTCTACGCCATCAACAGTGGTTACCAATCATGCGGGGCGCTCTTTGGGAAGC 3102
Db |||||
QY 510 CTGGTATCTCTACGCTATCAACAGTGGCTACCAAGTCAATGGGGTCCGCTCTTTGGGAAGC 451
Db |||||
QY 3103 TCTTCTTTCCTCTCTGGGTGATTTGTTCACTTATACCCATCTCTCAAGGCTCTTATGGCA 3162
Db |||||
QY 450 TCTTCTTTCCTCTCTGGGTGATTTGTTCACTGACCTGTACCCCTCTTGAAGGCTCTCATGGGTC 391
Db |||||
QY 3163 GGGAAAACCGCACACCGACGATTTGTCATGCTGGGTGTCCTCTCTGCTTCTATCTTCT 3222
Db |||||
QY 390 GGCAGAACCGTACACCGACCAATTTGTTGTTGGGCCATCTCTCTTGCATCGATCTTCT 331
Db |||||
QY 3223 CCTTGTCTGGGTTCGTTGTGATCCATTCATCCCGTCTCGCTGGGCCCAATATCCAA 3282
Db |||||
QY 330 CCTTGTCTGGGTTCGTTATCGATCCATTCACCCCGTGTACAGGGCCCTGATACCCAA 271
Db |||||
QY 3283 CCTGTGCTCAACTCTAGGAAAGTGGGAGTTTGTAGAGACAGAA 3328
Db |||||
QY 270 AATGTGTATCACTCTAGGAGTGAGCAAGATAGTAGGCAAGAA 225
Db |||||
```

Search completed: August 23, 2004, 11:47:51

Job time : 1540 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:15:51 ; Search time 8675 Seconds
(without alignments)
12481.875 Million cell updates/sec

Title: US-09-900-237A-29

Perfect score: 3626

Sequence: 1 gcacgaggaaccccgctoca.....tcacattttggaggagtttt 3626

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST: *
1: em estba: *
2: em esthum: *
3: em estin: *
4: em estcu: *
5: em estov: *
6: em estpl: *
7: em estro: *
8: em htc: *
9: gb est1: *
10: gb est2: *
11: gb htc: *
12: gb est3: *
13: gb est4: *
14: gb est5: *
15: em estfun: *
16: em estom: *
17: em gss hum: *
18: em gss inv: *
19: em gss pln: *
20: em gss vrt: *
21: em gss fun: *
22: em gss nam: *
23: em gss mus: *
24: em gss pro: *
25: em gss rod: *
26: em gss phg: *
27: em gss vrl: *
28: gb gss1: *
29: gb gss2: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2572.2	70.9	3763	11	AY108113	AY108113 Zea mays
2	2402	66.2	3728	11	AY112236	AY112236 Zea mays
3	2199	60.6	3696	11	AY110079	AY110079 Zea mays
4	1373.2	37.9	3898	11	AY110415	AY110415 Zea mays

5	1320.4	36.4	2872	11	AY104236	AY104236 Zea mays
6	1319.8	36.4	3788	11	AY103701	AY103701 Zea mays
7	1315.6	36.3	3783	11	AY104730	AY104730 Zea mays
8	1267.8	35.0	3897	11	AY103655	AY103655 Zea mays
9	882	24.3	1874	14	CD726831	EST027 Cu
10	762.2	21.0	880	12	BM816138	BM816138 HCl09B12
11	751.6	20.7	916	12	BM816139	BM816139 HCl05G09
12	750.4	20.7	762	13	BQ802778	BQ802778 WHE2829 H
13	729.4	20.1	870	12	BG368813	BG368813 HVSMB1002
14	697.8	19.2	723	14	CD930401	CD930401 GR45.111C
15	689.8	19.0	701	14	BQ578769	BQ578769 WHE0309 E
16	689.6	19.0	767	13	BQ282057	BQ282057 WHE03028 D
17	689.4	19.0	710	13	BQ282057	BQ282057 WHE03028 D
18	687.8	19.0	750	10	BF624748	BF624748 HVSME001
19	685	18.9	706	14	CD877344	CD877344 AZO4.100C
20	684.2	18.9	712	13	BQ620630	BQ620630 TaLR1138D
21	683.8	18.9	978	12	BG321258	BG321258 Zm04.03f1
22	682	18.8	702	14	CD894946	CD894946 G118.127I
23	670.4	18.5	865	14	CD440398	CD440398 ELOIN0554
24	669.4	18.5	718	14	CA501842	CA501842 WHE4039 D
25	668.6	18.4	755	10	BF259426	BF259426 HVSMEF001
26	668	18.4	694	14	CD880265	CD880265 F1.001D08
27	664.2	18.3	680	14	CD878619	CD878619 AZO4.103D
28	664.2	18.3	966	14	CA247044	CA247044 SCCCAM209
29	663	18.3	705	9	AJ433167	AJ433167 AJ433167
30	651.4	18.0	659	12	BQ214460	BQ214460 BU214460
31	650.4	17.9	702	14	CD907188	CD907188 G468.106D
32	649.8	17.9	874	14	CD433620	CD433620 ELOIN0313
33	649.6	17.9	678	9	AJ433236	AJ433236 AJ433236
34	648.6	17.9	882	14	CK157933	CK157933 FGAS03913
35	647.6	17.9	872	14	CK155617	CK155617 FGAS03845
36	644.4	17.8	659	13	BQ620294	BQ620294 TaLR1173A
37	638.6	17.6	905	13	CA159025	CA159025 SCEZR2312
38	637.6	17.6	676	13	BQ468579	BQ468579 HM01K12T
39	633.2	17.5	700	14	CB882206	CB882206 HL01D01w
40	631.4	17.4	813	14	CK199867	CK199867 FGAS00837
41	630	17.4	654	9	AV914933	AV914933 AV914933
42	628.4	17.3	661	12	BQ277746	BQ277746 BJ277746
43	628	17.3	656	13	BQ620013	BQ620013 TaLR1138D
44	625.4	17.2	671	14	CB868804	CB868804 HCO9H06w
45	625.2	17.2	646	14	CA500605	CA500605 WHE4022_A

ALIGNMENTS

RESULT 1	AY108113	AY108113	3763 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	AY108113	Zea mays	PCO126465	mRNA	sequence.	
DEFINITION	AY108113	Zea mays	PCO126465	mRNA	sequence.	
ACCESSION	AY108113	Zea mays	PCO126465	mRNA	sequence.	
VERSION	AY108113.1	GI:21211191				
KEYWORDS	HTC.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
REFERENCE	1	(bases 1 to 3763)				
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2	(bases 1 to 3763)				
AUTHORS	Coe,E.H.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the					


```
QY 1864 ATAGGAATGATCGATATGCAACAGGAAACACTGCTCTTTTGTATATTAATTAATGAGGGCC 1923
Db 2136 ACAGGAATGATCGATATGCAACAGGAAACACCGTGTCTTTTCGATATTAATTAATGAGAGGTC 2195
QY 1924 TTGACGGCATTTCAAGGACCAAGTTTATGTGGAACTGTTGTGTTTCAACAGAAACAGCTA 1983
Db 2196 TTGATGGCATCCAAAGACCAAGTTTATGTGCGAACTGGCTGTGTTTCAACCGAAGAGCTC 2255
QY 1984 TCTATGGTTATGAGCCGCCAAATTAAGCGGAAGAGCCAGGTTTCTTGGCATCACTATGTG 2043
Db 2256 TATATGGTTATGAGCCGCCAAATTAAGCGGAAGAGGTTGTTTCTTGTGTCATCACTATGTG 2315
QY 2044 GGGCAAGAGAGGAGCAAGCAAGTCAAGAAAGAGAGCTCAGATAGAAAGAAAGTCAACA 2103
Db 2316 GGGTAGGAAGAGGCAAGCAAGTCAAGAAAGAGAGTCAAGAAAGAGTCAAGAGAGTCAAG 2372
QY 2104 AGCATGTGACAGTCTCTGTTCAGTATTCATTCGAAGACATAGAGGAGGTTGTTGAAG 2163
Db 2373 AGCATGTGACAGTCTCTGTGCCAGTATTCACCTTGAAGATATAGAGGAGGAGTTGAAG 2432
QY 2164 GTGCTGGTTTGAATGAGAAATCAGTTCTCATGTCTCAATGTCTCAATGAGCTTAGAAGAGAT 2223
Db 2433 GGGCTGGATTTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCTTAGAAGAGAT 2492
QY 2224 TTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGTTGGTGTCTTCACT 2283
Db 2493 TTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGTTGGTGTCTTCACT 2552
QY 2284 CTTCACTCCAGAAATCTTTTGAAGAGCTATCCATGTCATAGTTTGTGGCTATGAG 2343
Db 2553 CCGCAACTCGGAGTCTCTTCAAGAGAGCTATCCATGTTATAGCTGTGGCTATGAG 2612
QY 2344 ACAAGTCTGAATGGGGAATCGAGATGTTGGTGTGATCTATGATCTGTCAGAGATATTC 2403
Db 2613 ACAAGACTGAATGGGGAATCGAGATCGGCTGATCTACGGTCTGTGACAGAGACATTC 2672
QY 2404 TTACTGGATTCAAGTACAGCAAGAGGCTGGGTTCACTCTATTCATGCTCCAGCGCC 2463
Db 2673 TCACCGGATTCAGAGATGCAAGCGGAGGCTGGGTTCACTCTATGCTGCTCCAGCGCC 2732
QY 2464 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTCAAGATCGTCTGAACCAAGTCTGCGGT 2523
Db 2733 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTGCAAGCTCTGACCAAGTCTGCGGT 2792
QY 2524 GGGCTCTGGTTCTGTTGAATCTTTTCAAGCGGATTCGCGGCTTATGATGCTAGC 2583
Db 2793 GGGCTCTGGTTCTGTTGAATCTTTTCAAGCGGATTCGCGGCTTATGATGCTAGC 2852
QY 2584 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCTAA 2643
Db 2853 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCTAA 2912
QY 2644 CTTCTCTCCGCTTCTAGTCTATTGATATTCGCTGCTATCTGCTCACTGCTCACTGAAAGT 2703
Db 2913 CTTCTCTCCGCTTCTAGTCTATTGATATTCGCTGCTATCTGCTCACTGCTCACTGAAAGT 2972
QY 2704 TCATCATGCCAGATTAGCAACTTGGCCAGATCTGTTGTTATGTTGCTTCTTCTTCAA 2763
Db 2973 TCATCATGCCAGATTAGCAACTTGGCCAGATCTGTTGTTATGTTGCTTCTTCTTCAA 3032
QY 2764 TTTTTCGCCACTCGTATCCTTGGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2823
Db 3033 TTTTTCGCCACTCGTATCCTTGGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3092
QY 2824 GGAATGAACAGTTCTGGGTCAATGGAGTATCTCTGCACATCTGTTTGGCGTCTTTCAGG 2883
Db 3093 GGAACGAGCAGTTCTGGGTCAATGGAGTATCTCTGCACATCTGTTTGGCGTCTTTCAGG 3152
QY 2884 GTCTTCTGAGTGTCTGCGGTATCGAACCAACTCTCTGCTCACTCTCAAGGCTTAATG 2943
Db 3153 GCTGCTCAAGGTGCTGGCGGATCGAACCAACTCTCTGCTCACTCTCAAGGCTTCTG 3212
```

```
QY 2944 ACGAAGAGCGGACCTTTGCTGAGCTCTACATGTTCAAGTGGAGCAGCTTCTCATCCCTC 3003
Db 3213 ACGAGACGGCGACTTCGCGGAGCTGTACATGTTCAAGTGGAGCAGCTCTCTGATCCGC 3272
QY 3004 CGAGGACCAATTTTATGATCAATTAACATGTTGGTGTGTTGCTGTCGACCTCTACGCCATCA 3063
Db 3273 CCACCACCAATCTGATCATCACTGTTGGGCTGTGTCGCGGCTCTCTACGCCATCA 3332
QY 3064 ACAGTGGTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGGTGA 3123
Db 3333 ACAGCGGATACAGTCTGTTGGGCGGCTCTCTCGCAAGCTCTTCTTGGCTTCTGGGTGA 3392
QY 3124 TTGTTCACTTATACCAATTCCTCAAGGTTCTTATGGGAGGCAAAACCGCACACCGACGA 3183
Db 3393 TGTTCACCTGTACCGGTTCTCAAGGGCTCATGGGAGGAGCAACCGCACACCGACCA 3452
QY 3184 TTGTCATCTCTGTTGGGCTGCTCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3243
Db 3453 TGTGCTGCTCTGTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3512
QY 3244 ATCCATTCATCTACCGTCTGCTGCGCCAAATATCCAACTGTGGCATCAACTGCTAGG 3303
Db 3513 ACCCTTCAACACCGCGTCACTGGCCGGATACCCAGACGTGTGGCATCAACTGCTAGG 3572
QY 3304 ABAAGTGGAGTTT 3316
Db 3573 GAAGTGAAGAGTT 3585
```

RESULT 2
AY112236
LOCUS AY112236 3728 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL1160_1 mRNA sequence.
ACCESSION AY112236
VERSION AY112236.1 GI:21216826
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 3728)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,R.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3728)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from zmdb and may be found by BLAST
searching at MSL, maizemap.org; Zmdb, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from Zmdb:
www.zmdb.iastate.edu.

FEATURES
Location/Qualifiers
1..3728
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"


```
QY 2136 CTCGAAGACATAGAGGAGGTGTTGAAGGTGCTGGTTCATGATGAGAAATCAGTTCTC 2195
Db 2311 NNNNNNNNTAGAGAGGAGTGGAGGCTGGATTGATGATGAGAAATCAGTTCTT 2370
QY 2196 ATGTCTCAAAATGAGCTTAGAGAGAGATTGGCCAGTCAGCAGCAATTTGTTCCTCCACT 2255
Db 2371 ATGTCTCAAAATGAGCTTGAGAAGAGATTGGCCAAATCTGCAGCTTTGTTCGCTCCACT 2430
QY 2256 CTGATGGAATATGGTGGTGTCTCTAGTCTCTCACTCCAGATCTCTTTTGAAGAAGCT 2315
Db 2431 CTGATGGAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2490
QY 2316 ATCCATGTCTAAGTTGTGGCTATGAGGCAAGCTCTGAATGGGAACTGAGATTGTGTG 2375
Db 2491 ATCCATGTCTAAGTTGTGGCTACGAGGCAAGATTGAATGGGAACTGAGATTGTGTG 2550
QY 2376 ATCTATGGAATCTGTCTACAGAAGATATTTCTTACTGGATTCAAGATGCACGCAAGAGCTG 2435
Db 2551 ATCTATGGAATCTGTACGGAAGATATTTCTACTGGTTCGAAGTCTCAAGATGCACGCAAGAGCTG 2610
QY 2436 CGTTCAGTCTATTGCAATGCCCAAGCGCCAGCTTTCAAGGATCTGCCGCCCATCAATCTT 2495
Db 2611 CGTTCAGTCTATTGCAATGCCCAAGCGCCAGCTTTCAAGGATCTGCCGCCCATCAATCTC 2670
QY 2496 TCAGATCGTCTGAACCAAGTCTCGGTGGGTCTCGTCTCTGTTGAAATCTTTTTCAGC 2555
Db 2671 TCAGACCGTCTGAACCAAGTCTCGGTGGGTCTCGTCTCTGTTGAAATCTTTTTCAGC 2730
QY 2556 CGGCATTGCGCCTTATGATGATACGAGGCGCCTCAAGTCTCTGGAGAGATTTCGT 2615
Db 2731 CGGCATTGNNNNNTATGATGATACGAGGACGCGCTCAAGTCTCTGGAGAGATTTCGCG 2790
QY 2616 TACATCAACACCAACATTTACCACCTAAGTCTCTCCGCTCTAGTCTATTGTATATG 2675
Db 2791 TACATCAACACCAACATTTACCACCTAAGTCTCTCCGCTCTAGTCTATTGTATATG 2850
QY 2676 CTTGTCTATCTGTCTCTACTGGAAGTTTCATCATGCGAGAGATTAGCAACTTGGCCAT 2735
Db 2851 CTTGCATCTGCTGCTACGCGGAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGT 2910
QY 2736 ATCTGTTTCAATTCGCTCTCTCTTCAATTTTTCGCACTGGTATCTTCGAGATGAGTGG 2795
Db 2911 ATCTGTTTCAATTCGCTCTCTCTTCAATTTTTCGCACTGGTATCTTCGAGATGAGTGG 2970
QY 2796 AGTGTGTGGGATTCAGGAGTGGTGGAGATGAACAGTTCTGGGTCTATTCGAGGATC 2855
Db 2971 AGCGGCGTGGGATTCAGGAGTGGTGGAGATGAACAGTTCTGGGTCTATTCGAGGATC 3030
QY 2856 TCTGCACATCTCTTTCGCTCTCTTTCAGGCTCTCTGAAGTCTCTTCGCGGTATCGACACC 2915
Db 3031 TCCGCCACCTCTTCGCGCTCTTCAGGCTCTCTTCAGGCTCTCTTCGCGGTATCGACACC 3090
QY 2916 AACTTCACTGTCACTTCAAGGCTAATGACGAAGGCGACTTTTCTGAGCTCTACATG 2975
Db 3091 AACTTCACTGTCACTTCAAGGCTAATGACGAAGGCGACTTTTCTGAGCTCTACATG 3150
QY 2976 TTTCAAGTGGAGAGCTTCTATCTCTCGAGGAGATGAACAGTTTCATCTAATGAGTGTGT 3035
Db 3151 TTTCAAGTGGAGAGCTTCTATCTCTCGAGGAGATGAACAGTTTCATCTAATGAGTGTGT 3210
QY 3036 GTCTGTGTGGCAGCTCTCTACGCCATCAACAGTGGTTTACCAATCATGCGGCGCTCTTT 3095
Db 3211 GTTGTTCGCGGATCTCTCTACGCCATCAACAGGAGTACAGTCTGAGGCTCTCTCTTCT 3270
QY 3096 GGAAGAGCTTTCTTTCGCTCTTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGCTCTT 3155
Db 3271 GGAAGAGCTTTCTTTCGCTCTTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGCTCTT 3330
QY 3156 ATGGGAGGCAAAACGGACACAGGAGATGATCTCTCTGGGCTGTCTCTCTCGTCTCT 3215
Db 3331 ATGGGAGGCAAAACGGACACAGGAGATGATCTCTCTGGGCTGTCTCTCTCGTCTCT 3390
QY 3216 ATCTTCTCTCTGTGGGTCTGTTGATCCATTTCACTTACCCTCTCGCTGGCCCCAAT 3275
```

```
Db 3391 ATCTTCTCTCTGTGGGTTCGATCGATCGCTTACCACCGCTCTCTGCCCCGAT 3450
QY 3276 ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGCA-GTTTGTAGAGACAGAAA 3329
Db 3451 ACTCGAAAGCTGTGGCATCAACTGCTAGGAGGTGGAAGGTTTGTAGAACAGAGA 3505

RESULT 3
AY110079
LOCUS Zea mays CL1164_1 mRNA sequence. 3696 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays
ACCESSION AY110079
VERSION AY110079.1 GI:21214162
SOURCE HTC.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3696)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
Location/Qualifiers
1..3696
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630055"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Query Match 60.6%; Score 2199; DB 11; Length 3696;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 2549; Conservative 0; Mismatches 681; Indels 9; Gaps 3;

QY 66 GCGACGCGGCGCGCCCTGAAGTCCGGGAGGACGCGGCGCGGAGCTGTGCCAGATCTGC 125
Db 257 GACGGCGGCGGCGCGCCAGCAATTCGGGAGAGCATGTGGCGGCGGAGCTGTGCCAGATCTGC 316
QY 126 GCGGACGCGCTGGGCGCCACCGTTGGACGGCGAGCTTTCAACCGCTCGAGAGCTTCGCCGC 185
Db 317 GCGGACGCGCTGGGCGCCACCGCGCGAGCGCGACCTCTTCCACCGCTCGAGCTTCGCCGC 376
QY 186 TTCCCGTCTCCCGCCCTGTACGACGACGAGCGGAGGAGGACCCAGGCTTCGCTC 245
Db 377 TTCCCGTGTTCGCGCCCATGTCTACGAGTACGAGCGCAAGGACGCGCGCTGCGCG 436
QY 246 CAGTCAAGACCAAGTACAGCGCCACAGAGGAGCGCGAGTCCGCGGCGGAGGAGG 305
Db 437 CAGTCAAGACTTAGTACAGCGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 496
```

Qy	306	GACGACACTGATCCGATGATGTAGTGTCAATACCTACCTGSCATGSGAGGAC	365	Qy	1386	GACAAAGTCAGGCTTCATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAGAA	1445
Db	497	GAGGATGTGATGCTGACGATGTGAGTGACTACAATCACCAAGCATCTGGCAACACAGGAT	556	Db	1571	GACAAAGTTCAAACCTCATTTGTGAAGAAACGCGGNNNNNNNNNNNNNNNNNNNNNNNN	1630
Qy	366	CAGAAGCAGAAAGATTGCTGACAGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTGGC	425	Qy	1446	TTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAGTCTCCCGAGGAGGATGG	1505
Db	557	CAGAAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGCAAACTCACGTGGCAGTG--	614	Db	1631	NNCAAAGTTGCTATCAATGCTCTGTAGCCAAAGGCACNNNNNNNNNNNNNNNNNNNNN	1690
Qy	426	AATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAAGTATGACAGTGA	485	Qy	1506	ATCATCAGATGGCACACCATGGCCAGGAAACAAATACAGGAGGATCATCTCTGGAATGATT	1565
Db	615	-AATTTGGCCTGGCTGAAGTATGACAGCGGTGAATTTGGGCATGGGAAGTATGACAGTGT	673	Db	1691	ATCATGCAAGATGGTACACCTTGGCTTGGGAAACAATCTAGGAGCAATCTCTGGAATGATT	1750
Qy	486	GAGATCCCTAGGAGTACGTCCTTCACTCAACCAAGCCAGATGTCAGGAGAAATCCCT	545	Qy	1566	CAGGTTTCCCTCGTCCACAGTGGTGGCTTGTATCTAGGAGTAAATGAGTCTCCCGGTTTA	1625
Db	674	GAGATCCCTCGTGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCT	733	Db	1751	CAGGTTTCCCTGGGTACAGTGGAGGGCTTGAAGTTGAAGCAATGAATCTCTCGTTG	1810
Qy	546	GGAGCTTCGCTCATCATCATGATGTCCTTACGGGGAACATCAGCAACGTCGTCCTCG	605	Qy	1626	GTTTATGTGCTCGTGAAAGCGCTCTGGGTTCCAGCACCAAGAGGCTGGTGGCATG	1685
Db	734	GGAGCTTCCCTCATCATATGATGTCCTGTTGGGAACATTGGCAGGCGTGGACATCAA	793	Db	1811	GTTTATGTGCTCGTGNNNNNCGTCTGGATTCGAATCACAAGAGGCTGGTGGCATG	1870
Qy	606	TTTCCCTATGTGAATCATTCACCAATTCGTCAGGAGTTCCTCCGCGATATTGGGAAT	665	Qy	1686	AATGCCCTTGTCTCGTGTCTCAGCTGCTTACTTAATGGAACAATACATGTTGAATCTTGAT	1745
Db	794	TTTCCCTATGTAATCATCTCCAAACCCATCGAGGGAGTTCTCCGCTAGCCCTTGGCAAT	853	Db	1871	AATGCACTTGTTCGTGTATCAGCTGCTTACTTAATGGGCAATACATGTTGAATCTTGAT	1930
Qy	666	GTTGCTCGGAAAGAGAGATTGATGGCTGGAATGAAGCAGGACAAAGGCTGGATTCCT	725	Qy	1746	TGTGATCACTACATCAACAACAGCAAGGCTGTCCGAGAGCTATGTGCTTCCTAATGGAT	1805
Db	854	GTTGCAATGGAAGAGAGGTTGATGGCTGGAATGAAGTGAATGA--AGTAAAGGTGCAATTCCT	910	Db	1931	TGTGACCACTACATCAATATAGCAAGGCTCTTCGAGAGCTATGTGCTTCCTTATGGAC	1990
Qy	726	ATGACTAATGGGCAAGCAATGTCTCCCTGAAAGTCCGGCAGCTACTGATCGATGCA	785	Qy	1806	CCAAACCTAGTCCGCAAGTCTGTTATGTGCAAGTCTCCCAAAAGGTTGATGGATGAT	1865
Db	911	ATGACCAATGGAACAACCAATGTCTCCATCAGAGGGCGTGGATGATGATGCT	970	Db	1991	CCAAACCTAGGAGGAATGTCTGTTATGTCCAATTTCCCTCAGAGGTTGATGGTATGAT	2050
Qy	786	TCTACTGAATACAACTGGAAGCGCTTACTGATGATGAATCGAATCCGAGCCCTCTATCT	845	Qy	1866	AGGAATGATCGATATGCAACAGCAACACTGTCTTTTTCGATATTAACCTTGAGGGGCTT	1925
Db	971	TCTACTGATTATACATGGAAGATGGCTTACTGAATGATGAATCGGCAACCTCTATCT	1030	Db	2051	AGGAATGACCGATATGCAACAGCAACACTGTGNNNNNCGATATTAACCTTGAGAGGCTT	2110
Qy	846	AGAAAGTCCCATTTGCTCTCCAAATAAATCCCTACAGAAATGTCATTTCTCGCG	905	Qy	1926	GACGGCAATCAAGAACAGTTTATGTGGAACTGGTTGTGTTTTCACAGAACAGCTATC	1985
Db	1031	AGAAAGTGCAATTTCTTATCCAGAAATTAATCCGTACAGATGGTCATTTGCTACGT	1090	Db	2111	GACGGCAATCAAGGGCCAGTTTATGTGGAACTGGTTGTGTTTTCACAGAACGGCCCTA	2170
Qy	906	TTGGTTGTTCTTAAGCATCTTCTGSCACTACCGTCTCAAAATCCTGTGCGTAAATGCAATC	965	Qy	1986	TATGGTTATAGCCCCCAATTTAGGCGAAGAACGAGGTTTCTTGGCATCACTATGTGGG	2045
Db	1091	TTGGCTGTTCTATGCAATTTCTTCGCTACCGTATCACATCTCGTGAACAATGCATAT	1150	Db	2171	TATGGTTATAGGCTCCAGTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2230
Qy	966	CACTGTGGCTTTTATCTGTTATATGAGATTTGGTTTTCCTGATGACTGAT	1025	Qy	2046	GGCAAGAAAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAAAGTTCGAAACAAG	2105
Db	1151	CCACTGTGGCTTTTATCCGTCATATGTGAGATCTGGTTTGTCTTGTCTGATTTTGGAT	1210	Db	2231	NNAAGNNNNAGCTCAAAATCTAAGAA--GAGCTCGGAAGAAGAGTCAATAGA	2287
Qy	1026	CAGTTCCGGAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGATAGACTGGCTTTAAGG	1085	Qy	2106	CATGTGGACAGTTCTGTTCAGTATTCATCTCGAAGACATAGAGGAGGTTGTTGAAGGT	2165
Db	1211	CAGTTCCCAAGTGGTCCCNATCAACCGTGAACATACCTTGATAGACTGGCTTTAAGG	1270	Db	2288	CACGACACAGTTCTGTACCAAGTATTAAATCTCGAAGATATAGAGGAAGGATTTGAAGGT	2347
Qy	1086	TATGACCGAGAAGTGAACCGTCTCAGTTGGCTGCTGTGACATPATTTGTGACGACATC	1145	Qy	2166	GCTGGGTTTATGATGAGAAATCAGTTCTCATCTCTCAAAATGAGCTTGTGATGAGAGATTT	2225
Db	1271	TATGACCGAGAAGTGAACCAATCTCAATTAGCTCCTGTGATANNNNNGTCAGTACTGTG	1330	Db	2348	TCTCAGTTTATGATGAGAAATCGCTGATTTATGTCTCAAAATGAGCTTGGAGAAGATTT	2407
Qy	1146	GACCCCTTGAGAGGACCACTTATCGTCACTGCCAAGCTGTGCTATCCATCTTGTCTGTT	1205	Qy	2226	GGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGTGTTCTTCAGTCC	2285
Db	1331	GATCCAATGAGGAGGCTCTCTTGTCACTGCAAAATACGTGCTTCCATCTTGTCTGTC	1390	Db	2408	GGCCAGTCCAGTGNNNNNAGCTCTACTCTGATGGAATATGGTGTGTTCCCAATCT	2467
Qy	1206	GATTATCCGTTGGAAGGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT	1265	Qy	2286	TCCACTCCAGAAATCTCTTTTGAAGAGAGCTATCCATGCTCATTAAGTGTGGCTATGAGAC	2345
Db	1391	GATTATCCGTTGACAGGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT	1450	Db	2468	GCAACTCENN	2527
Qy	1266	TTTGACGCAATGGCTGAGATTCAGATTTGCTAGGAAATGGGTAACATTTGTGGAAGAG	1325	Qy	2346	AAGTCTCAATGGGAACTGAGATGGTTGGATCTTATGATCTGATCAAGAGATATTCTT	2405
Db	1451	TTTGATGCTCTCTCTGAAATTCAGAGTTTGTGAGAAATGGGTTCCGTTCTGTAAGAAG	1510	Db	2528	AAAACTGCTGGGGAATGAGATGGGTGGATCTATGTTCTGTTTCAAGAGACATCTC	2587
Qy	1326	TATGACATGAAACCGAGCTCCGAGTTTACTTTTCCAGAGAAATGATTACTGAAA	1385	Qy	2406	ACTGGATTCAAGATGACGCAAGAGGCTGGGCTTCACTTCTATTGCTATGCCAACGCGCCA	2465
Db	1511	TACAACATAGAGCCTANGGCCCCGGAATGGTACTTTGCTCGAAGAAATGATTACTTGAA	1570	Db	2588	ACCGNN	2647
				Qy	2466	GCTTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTCGGTGG	2525

Db	2648	GCTTTCAAGGGATCTGCTCCATTCAACCTTTTCGATCGTTTGAATCAAGTCTCTTCGGTGG	2707
Qy	2526	GCTCTCGGTTCTGTTGAAATCTTTTCAGCGCGCATTCGCCCTTATGTGTATGGTACGGTA	2585
Db	2708	GCTCTTGGTCCATTGAAATCTTTTCAGCAGGCATTTGCCATATGATGGTATGGCTATGGA	2767
Qy	2586	GGGCGCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTATTACCCACTAACC	2645
Db	2768	GGCGGCTTAAATTCCTGGAGAGATTTGCTTATATCAACAACAATTTATCCACTACA	2827
Qy	2646	TCTCTCCGCTTCTAGTCTATTGTATATTTGCTCTGCTATCTGTCTGCTCACTGGAAGTTC	2705
Db	2828	TCAATCCGCTCCCTCGTACTGTCATATTCGCACAGTTTGTCTCTCACTGGGAAGTTC	2887
Qy	2706	ATCATGCGAGAGATTAGCAACTTCGGCCAGTATCTGGTTTCATTGTGCGTCTCTTCTCAATT	2765
Db	2888	ATCATCCCAAGATTAGTAACCTAGAGAGTTTGGTTTATATCGCTCTTTATCTCAATC	2947
Qy	2766	TTCCGCACGTGATCTCTTGAGATGAGGTGGAGTGGTGTGGGCATTGACGAGTGGTGGAGG	2825
Db	2948	TTTGCCACTGGTATCTCTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	3007
Qy	2826	AATGAA CAGTCTCGGTCATTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTCAGGGT	2885
Db	3008	AACGAGCAGTCTCGGTCATTGGTGGTATTTCTGCGCATTTATTTTGCGCTCTTCCAGGT	3067
Qy	2886	CTTCTGAAGGTCTTTCGGGTATCGACACCAACTTCACCTGCTCACTCAAGAGGCTAATGAC	2945
Db	3068	CTCTGAAGGTCTTTCGTGGTATCGACAGAGCTTCACTGTCACTTAAAGGCACTGAC	3127
Qy	2946	GAAGAAGCGCACTTTCGTGAGCTCTACATGTTCAAAGTGACGACGCTTCTCATCCCTCCG	3005
Db	3128	GAAGAAGTGATTTTTCGGAGCTCTACATGTTCAAAGTGACGACGCTTCTGATCCCACTCA	3187
Qy	3006	ACGACCAATTTTGATCATTAACATGTTGGTGTGCTTGGCACTCTACGCGCATCAAC	3065
Db	3188	ACCACTATTTTGATCATCAACTGCTGGCGGTGGTGGCACTCTTCTTGGCTTCTGGGTGAT	3247
Qy	3066	AGTGGTTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGGTGAT	3125
Db	3248	AGCGGTTACAGTATGGGGAACCTTTTCGGGAAGCTCTTCTTGGGTCTGGGTGAT	3307
Qy	3126	GTTCACTTATACCAATTCCTCAAGGTCCTTATGGCAGGCAAAACCGCACACGACGAT	3185
Db	3308	GTCCACCTGTACCCCTTCTCAAGGCGCTCATGGGAGCAGAACCGCAGCCACAT	3367
Qy	3186	GTCACTGTCTGGGCTGCTCTCTCGTTCATCTTCTCTGTGTGGGTTCGTGTGAT	3245
Db	3368	GTCTGTCTGGGCTATCTCTTGGTGCATCTTTTCCCTGATGTGGTTCGTATCGAT	3427
Qy	3246	CCATTCACTACCGTCTCGCTGGCCCAATATCCAAACCTGGGATCACTGCTAGGA	3304
Db	3428	CCATTCAACCCGGGTACTGTCTCTGATATCGGAAATGTGGCATCACTGCTAGGA	3486
RESULT 4			
AY110415			
LOCUS	AY110415	3898 bp	linear
DEFINITION	Zea mays CL1166.1	mrna	HTC 17-OCT-2002
ACCESSION	AY110415		
VERSION	AY110415.1		
KEYWORDS	HTC		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
TITLE	1 (bases 1 to 3898)		
	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		


```

Db      2917 GCATTGAAGATTGGTGAGAAATGAGCAGTTTGGGTTATTGGTGCCACCTCTGCCAATC 2976
QY      2866 TGTTCGCCGCTCTTTCAGGGTCTTCTGAAGTGCTTGCCGTAATCGACACCACTTCACTG 2925
Db      2977 TCTTCGAGTGTCCAGGGTCTCTGAAAGTGTGGCTGGGATGATACCACTTCACAG 3036
QY      2926 TCACCTCAAAGGCTAATAGCAGAAAGGCGACTTTGCTGAGCTCTACATGTTCAAGTGGA 2985
Db      3037 TTACCTCAAAGGCATCTGATGAGGATGGGACITTTGCTGAGCTATATGTGTTCAAGTGA 3096
QY      2986 CGAGCGTCTCTATCCCTCCGACGACCATTTTGAATCAATTAACATGTTGGTGTGCTG 3045
Db      3097 CAGTTTGCTCAATCTCTCCGACCATGTTCTTGTCAATTAACCTGGTCGGAATGGTGCGAG 3156
QY      3046 GCACCTCTACGCCATCAACAGTGGTTACCAATCATGGGGCGCTCTTGGGAAGCTCT 3105
Db      3157 GAATTTGGTATGCCATTACAGTGGCTACCAATCTGGGGTCGCTCTTTGGAAGCTGT 3216
QY      3106 TCTTTGCCCTTCTGGGTGATGTTCACTATACCATCTCTCAAGGCTCTTATGGGCGAGC 3165
Db      3217 TCTTCTCGATCTGGGTGNTCTCCATCTGTACACGTTCTCAAGGCTCTCATGGGAAGC 3276
QY      3166 AAAACGGCACCGAGATTGTTCATCGTCTGGGCTGCTCTCGCTTCTATCTTCTCT 3225
Db      3277 AGAACCGCACACAAACATCGTCAATGCTGTGCTGCCATCCTCTTGCACTATCTTCTCT 3336
QY      3226 TGCTGTGGGTCTGTGTGATCCATTCACTACCC 3258
Db      3337 TGCTGTGGGTGAAGATCGATCTTTCATCTCCC 3369

RESULT 5
AY104236
LOCUS      Zea mays PC0121439 mRNA sequence.
DEFINITION
ACCESSION AY104236
VERSION   AY104236.1 GI:21207314
KEYWORDS  HTC.
SOURCE    Zea mays
ORGANISM  Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2872)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2872)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1. .2872
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637064"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the

```

overgo addressing of BACs in conjunction with the Maize Mapping Project"

Query Match 36.4%; Score 1320.4; DB 11; Length 2872;
Best Local Similarity 72.2%; Pred. No. 6.2e-265; Indels 10; Gaps 3;
Matches 1760; Conservative 0; Mismatches 666;

```

QY      829 CTGCGCAGGCTCTATCTAGAAAAGTCCCATTTGCTTCTCCCAAATAAATCCCTACAGAA 888
Db      14 CACGGCTACTCTAAGTCGATAGTTCCGATATCTCCAAACGAGCTTAACCTTTATCGGA 73
QY      889 TGTCTATCTCTGCGGTTGGTTCTTAAGCATCTTCTGCACTACCGTCTCACAATC 948
Db      74 TGTGATTTGTTCTCGGCTTATCATCTATGTTTCTTCTTCAATATCGTAATCACTCATC 133
QY      949 CTGTGCGTAATGCATACCCACTGTGGCTTTTATCTGTATATGTGAGATTGTTGGTTGCTT 1008
Db      134 CAGTGGAGATGCTTATGGGTTGTGGCTTGAATCTGTTATTTGTGAAGTTTGGTTGGCT 193
QY      1009 TATCTGGATATCTGGATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGGAGACCTACTTGG 1068
Db      194 TGTCTTGGCTTCTAGATCAGTTCCCAAAGTGTGATCTTATCAACCGTGAACCTTACTCTG 253
QY      1069 ATAGACTGGCTTTAAGGTATGACCGAGAAGGTGAACCGTCTCAGTT-GGCTGCTGTGTGAC 1127
Db      254 ATAGACTTGCATTTGAGATGATAGGAGGGGTGAGCCATCCAGTTGGGCTCCAATCGAT 313
QY      1128 ATATTGTGCAGTACAGTCGACCCCTTGAAGGAGGCACCTATCGTCACTGCCAACACTGTG 1187
Db      314 GTCTTTGTAGTACAGTGGATCCACTTAAGNAACCTCTCTAATTACTGGCAACTGTG 373
QY      1188 CTATCCATCTTGTGTTGATTATCCCGTGGACAAGTCTCTTGTGTATGTATCTCATATGAC 1247
Db      374 CTGTCCATCTTGTGCTGGATTACCCCTGTGTGACAAAGTATCATGTTAATGTTTCTGATGAC 433
QY      1248 GGAGCTTCAATGCTGACCTTTTGAACCATTTGGCCATTTGGCTGAGACTTCAAGTTTGTAGGAAATGG 1307
Db      434 GGTTCAGCTATGTTGACTTTTGAAGCGCTATCTGAAACCGCAGAGTTTTCGAAGGAAATGG 493
QY      1308 GTACCATTTGTGAAGAAGTATGACATTGAACCCAGAGTCCCGAGTCTTCTTTTGGCCAG 1367
Db      494 GTTCCCTTTTCGACGAACAACAATATGAACCTAGGGCTCCAGAGTTTACTTTTCTCTCGA 553
QY      1368 AAAATTGATTACCTGAAAGACAAAGTCCAGCCTTCATTGTTTAAAGACCGCGGGCCCATG 1427
Db      554 AAGATAGATTACCTAAAGGACAAAATACAACCTTCTTTTGTGAAGAAGGCGGCTATG 613
QY      1428 AAGAGAGATATGAAGAATTTAAATCAGGATTAATGCCCTTAGTCTTCTAAGGCAATGAAA 1487
Db      614 AAGAGGGAGTGTGAAGAGTTTCAAAGTACGGATCGATGCCCTTGTTCAAAAGCGCAAAA 673
QY      1488 GTCCCCCAGGGAAGGATGATCATCAAGATGSCACACCATCGCCAGGAAACAATACCAAG 1547
Db      674 ATACTGAGGAGGCTGGAACATGGCTGATGGCACTCTTGTGGCTTGGGAAATAACCTTGA 733
QY      1548 GATCATCTCGAATGATTCAAGTTTTCCTTGGTCAACAGTGGTGGCTTGTACTACTGAGGTT 1607
Db      734 GATCATCGMGAATGATCCAAAGTATTTCTTGGGCCACAGTGTGGGCTTGACACCGATGGG 793
QY      1608 AATGAGCTCCCCGTTTGTAGTTTATGTTCTGCTGAAAAAGCGTCTCGGGTTCACGACCAAC 1667
Db      794 AATGAGTTGCCACGGCTTGTATTGTTTCTGTTGAAAAAGAGGCGCAGGCTTCCAGCACCA 853
QY      1668 AAGAAGGCTGTGTCATGAATGCCCTTGTTCGTGCTCTCAGCTGCTCTTACTTAATGGCAA 1727
Db      854 AAGNAGGCTGTGTCATGAATGCTTTGATTGCGGTATCAGCTGTCTTCGACGAATGTTGCT 913
QY      1728 TACATGTTGAATCTTGAATGTGATCACTACATCAACAAACAGCAGGCTGTCCGAGAAAGCT 1787
Db      914 TATCTTCTTAATGTGGATTGTGATCACTACTTCAATAGACGAAGAGCTCTTAGAGAGCT 973
QY      1788 ATGTGCTTCTTAATGGATCCAAACCTAGGTCCGCAAGTCTGTTATTGTGAGTCTCCCAAA 1847

```


Db	2048	CTCTTTGTGTGTTCCAAAGTCTCTTAAAGGTGTAGCAGGATCGACACAAACTTTCACG	2107		
QY	2925	GTCACTCAAAGGCTAATGACGAAGAGCGCACTTTGTCTGAGCTCTACATGTTCAAGTGG	2984		
Db	2108	GTCAATCAAAGGCAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2167		
QY	2985	ACGACGCTTCTCATCTCCGACGACCAATTTTGATCAATTAACATGTTGGTGGTGGTGGT	3044		
Db	2168	ACAACCTTCTGATCCCCCCCCACCTGTGCTTGTGATTAACCTGGTTGGTATAGTGGCT	2227		
QY	3045	GGCACTCTACGCCATCAACAGTGGTTACCAATCATGGGGCGCGCTCTTTGGGAAGCTC	3104		
Db	2228	GGAGTGTGATGCTATCAACAGTGGCTACCAATCATGGGTCCCACTATTTCGGGAAGCTG	2287		
QY	3105	TTCTTTGCCCTTCTGGGTGATTTGTTCACTTATACCAATCTCAAGGCTCTTATGGGAGG	3164		
Db	2288	TTCTTTGCAATCTGGGTGATCTCCACCTCTACCTTCTCGAAGGCTCTCATGGGGAAG	2347		
QY	3165	CAAAACCGCACACCGGACGATTTGTCATCGTCTGGGCTGCTCTCGCTTCTATCTTCTCC	3224		
Db	2348	CAGAACCGCACACCGGACCACTCGTCAATGTTGGTCCGTCCTTCTGCTTCCATTTCTCG	2407		
QY	3225	TTGCTGTGGGTTGCTGTGATCCATTCACCTACCCGT	3260		
Db	2408	CTGCTGTGGGTGAAGATCGACCCCTTCATATCCCT	2443		
LOCUS	AY103701	3788 bp	mRNA	linear	HTC 16-OCT-2002
DEFINITION	Zea mays PC0120363 mRNA sequence.				
ACCESSION	AY103701				
VERSION	AY103701.1 GI:21206779				
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	1 (bases 1 to 3788)				
AUTHORS	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 3788)				
AUTHORS	Coe, E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				
FEATURES	Location/Qualifiers				
source	1..3788				
	/organism="Zea mays"				
	/mol_type="mRNA"				
	/db_xref="MaizeDB:637005"				
	/db_xref="taxon:4577"				
	/clone_lib="Maize Mapping Project/DuPont Consensus Library"				
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"				
ORIGIN					

Query Match	36.4%;	Score 1319.8;	DB 11;	Length 3788;
Best Local Similarity	65.3%;	Pred. No. 8.9e-265;		
Matches 2129;	Conservative	0;	Mismatches 1030;	Indels 103;
Gaps	9			
QY	89	CGGGAGGACACGGGGCCGGGACGTGTGCAGATCTGCCCGCTTCCCGGTCTGCGCGCCTGGGACACACGTT	148	
DB	296	CGCGCGCCCGGGCGAGGCGCGTGTCCAGATATGCGCGACAGAGTTCGGGGTGGGCTT	355	
QY	149	GGACGGCAGCGTCTTACCGCTGTGGAAGTCTGCCCGCTTCCCGGTCTGCGCGCCTGTCTA	208	
DB	356	CGACGGGAGCCCTTCGTGGGTGCAACGAGTGGCCCTTCCCGTCTGCGCGCCTGTCTA	415	
QY	209	CGAGCACAGCGCAAGGAGGACCCAGCGCTGCTCCAGTGTCAAGACAAGTCAACGCG	268	
DB	416	CGAGTACAGAGCGCCGAGGGTCTCAGCGCTGCCCGCAGTGCAGAACCCGCTACAAGCG	475	
QY	269	CCACAGAGGGACCCAGACGATCCCGCGGAGGAGGCGACACACTGATGCGGATGATGG	328	
DB	476	CCTCAAGGTTTGCCCGCGGTGGCTGGGGAACGAGGAGGACGCGCTGCAGCACTTGA	533	
QY	329	TAGTGACTTCAAC---TACCCTTGACTCTGGCACTGAGGACACAGAGCAAGATGTCTGA	385	
DB	536	GGGCGAGTTTGGCTCTCAGGACGGCGCCACGAGGACGACCCGAGTACGTGC CGA	595	
QY	386	CAGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTGGCAATGTGGCCACCCCAAGTA	445	
DB	596	GTCCATGTGCGGGCGCAGATGAGTACGCGCGCGCGCGACGCGCACCCGGACTTCAA	655	
QY	446	TGACAGTGGCAGATCGGCTCTCCAACTATGACAGTGGAGAGATCCTTAGGGATACGT	505	
DB	656	CCCCTGCCAAGTGC CGCTCTCCCAACG-----GCCAGATGGTTGATG	702	
QY	506	CCCTTCAGTCAACACGCCAGATGTTCAGAGAAATCCCTGGAGCTTCGCTGATCATCA	565	
DB	703	ACATCCCGCGAGCAGCACCGCTGTGCGCTCTATATGGGCGGCGCGCGCGCGCA	762	
QY	566	CATGATGTCCCTTAGGGGAACATAGACAGCGTCTCCGTTTCCCTATGTGAATCATTC	625	
DB	763	AGAGGATCCACCGCTCCCTTCGTGATTCACACCTTCCAGTGCACACGAGATCCATGG	822	
QY	626	ACCAATCCGTCAGGGAGTTCCTCGGCAGTATTGGGAATGTTGCTTGGAAAGAGAGT	685	
DB	823	ACCCGTCAAGGA--TCTGGCGCTTACGGATATGGCAGCGTGGCCCTTGGAAAGGAGAAT	880	
QY	686	TGATGGCTGAAAAATGAAGCAGGACAAAGGTCGCAATCCCATGACTAATGGGACAAGCAT	745	
DB	881	GGAGGCTGGAAGCAGAAAGCAGGAGCGCTGCAGCATGTTCAGGACGAGG-----	931	
QY	746	TGCTCCCTCTGAGGTGGGAGCTACTGACATCGATCTACTGAATACAACTGGA	805	
DB	932	-----TGGCGGTGATTGGGATGGGACGATGC	958	
QY	806	AGACGCTTTACTGAATGATGAATCGCCAGCCTCTATCTAGAAAAGTCCCAATGCTTC	865	
DB	959	AGATCTGCCACTAATGGATGAGCTAGGACGCCATGTCCAGAAAAAGTCCCTATATCATC	1018	
QY	866	CTCCAAATAAATCCCTACAGAATGGTCAATTGTTCTCGCGTTGGTCTCTTAAGCATCTT	925	
DB	1019	AAGCCGAATTAATCCCTACAGGATGATATCGTTATCCGTTGGTGGTTTGGGTTTCTT	1078	
QY	926	CCTGCACTACGCTCAAAATCCTGTGCGTAATGCAATCCCACTGTGGCTTTTATCTGT	985	
DB	1079	CTTCCACTACCGAGTATGATCCGGGAAAGATGCATTTGCAATGTGGCTCATATCTGT	1138	
QY	986	TATATGTGAGATTGGTTTCTTATCTCTGGATCTGGATCAGTTCCTCGAAGTGGTTTCC	1045	
DB	1139	AATCTGGAATCTGGTTTGCATGTCTCGATCTCTGATCTTGTGATCTCCAAAGTGGCTTCC	1198	
QY	1046	AATCAACCGGGACCTACCTTGATGACTGGCTTTAAGGATGACCGAAGGTGAAC	1105	
DB	1199	AATCGAGAGAGACTTACCTGACCGTTTGTCACTAAGGTTTGAAGAAGGTCAACC	1258	
QY	1106	GTCTCAGTTGGCTGCTGTGACATATTGTGACATGAGTGCACCCCTTGAAGAGCCACC	1165	

Db	1259	CTCTCAGCTGCTCCATCGACTCTTTGTGTCAGTACGGTTGATCCCAAGGAACCTCC	1318
Qy	1166	TATCGTCACTCGCAACACTGTGCTATCCATTTCTGTGTTGATATATCCCGTGGCAAGGT	1225
Db	1319	CTTGCTCACAGCGAACACTGTCTCTTCCATCTCTCTGTGATTATCCGGTTCGGAAGGT	1378
Qy	1256	CTCTTGCTATGTATCTGATGACGGAGCTTCAAAGCTGACTTTTGAAGCATTTGCTGTGAC	1285
Db	1379	CTCCGTGCTATGTTTCTGATGATGGTGTGCAATGCTTACGTTTGAAGCATTTGTCTGAAC	1438
Qy	1286	TTCCAGAGTTTGTCTAGGAAATGGGTACCATTTGTGAAGAAGTATGACATTGAAGCCAGAGC	1345
Db	1439	ATCTGAATTTGCAAGAATAAGGTTCTTTTCAGCAAAAGTTTAAATTCAGGCTCGTGC	1498
Qy	1346	TCCCGAGTTTACTTTTTCGCAGAAAATGTGATACCTTGAAAGACAAAAGTCAGCCTTCATT	1405
Db	1499	TCCTGAGTGGTACTTCCAAACAGAAGATAGACTACCTGAAAGACAAGGTTGCTGCTTCATT	1558
Qy	1406	TGTTAAAGACCGCGGGCCATGAGGAGAGAAATAGAGATTTAAATCAGGATAAATGC	1465
Db	1559	TGTTAGGGAGAGAGGGCGATGAAGAGAGAAATACGAGGAATTCAGGTAAAGTACATGTC	1618
Qy	1466	CCTAGTTTCTAAGGCAATTGAAGTCCCGAGGAAGATGGATCATGCAAGATGGCACACC	1525
Db	1619	CTTGTTTGCAAAAGCCCCAAAAGGTTCTGAGGAAGATGGAAGTCAAGATGGAAAGCCC	1678
Qy	1526	ATGCCACAGGAAAACAATACCAAGGATCATCTGGAATGATTCAGGTTTCTCTGGTCACAG	1585
Db	1679	CTGGCCTGGAAAACAACGTACGCGATCATCTGGAATGATTCAGGTATCTCTTGGCCAAAG	1738
Qy	1586	TGGTGGCCTTGATACTCAGGGTAATGAGTCCCGCTTTAGTTTATGTGTCGTGAAAA	1645
Db	1739	TGGCGGTGCTGATGTGAAGGAAATGAGTGGCCTCGCTGGTTTATGTCGAGAGAAAA	1798
Qy	1646	CGGTCTCTGGGTTCCAGCACACAAGAGGCTGTGCCATGAATGCCCTTGTTCGTGTCCTC	1705
Db	1799	GAGCCAGGTTTAAACCATCACAAGAAGCTGTGTCATGAATGCACTGGTCCGTGTCCTC	1858
Qy	1706	AGTGTCTCTTACTAATGGACAATACATGTTGTAATCTTGATTTGTGATCACTACATCAACAA	1765
Db	1859	TGCTGTCTTATCAATGCTGCATAACCTATTGAACTTTGGACTGTGATCACTACATCAACAA	1918
Qy	1766	CAGCAGGCTGTCGGAGAGCTATGTGCTCTCTAATGGATCCAAACCTTAGTCCGGAAGT	1825
Db	1919	TAGCAAGCCCATAAAAGAGGCTATGTGTTTCAATGATGGATCGTCTGGTGGGGAAGAAAT	1978
Qy	1826	CTGTATTGTGCAGTTCCTCCACAAGGTTTCATGGGATTGATAGGAATGATGCATATGCAAA	1885
Db	1979	GTGCTATGTACAGTTCCCTCAGAGTTTCAATGGTATTCGAAAAATGATCGATACGCTAA	2038
Qy	1886	CAGGAACATGTCTTTTGTATTAATCTTGAGGGCCCTTGACGGCATTCAAGAGCCAGT	1945
Db	2039	CAGGAACGTTGTCTTTTGTGACATCAACATGAAGGTTTGGACGGTATTCAGGACCCAT	2098
Qy	1946	TTATGTGGGAACTGTTGTGTTTTCAAGAAACAGGTATCTATGTTTATGAGCCCCCAAT	2005
Db	2099	TTATGTGGGTACTGGAATGTGTTTTCAGACGGCAGGCACTGTATGTTTATGATGTCCTTAA	2158
Qy	2006	TAAGGCGAAGAAAGCCA-----GGTTTCTGGGCTCACT	2038
Db	2159	AACGAAGAGCCACCATCAAGAACTTGCACTGTGGCCCAAGTGGTGCTCTTGTGCTG	2218
Qy	2039	ATGTGGGGCAAGAAGAGGCAAGCTCAAGAAAAAGGAGCTCAGATAAGAAAAA----	2095
Db	2219	CTGCACAGGAACAAGATAAAAAAGAACACTACAAAAACCAAGCGGAGAGAGAAAAAG	2278
Qy	2096	---GTCGAAACAGCATGTGGACAGTTCTGTGTCAGATTTCAATCTCGAAGACATAGAGGA	2152
Db	2279	ATTATTTTTCAGAGAAGCAGAAAAACCCATCTCCTGCATATGCTTTGGGTGAAATTGATGA	2338
Qy	2153	GGGTGTTGAAGGTGCTGGGTTTGTGATGATGAGAAATCAGTTCTCATGTCTCAATATGAGCTT	2212

```
Db 2339 AGGTGCTCCAGTGC-----TGATATCGAGAAGGCCGAATCGTAAATCAACAGAAACT 2392
QY 2213 AGAAGAAGATTGGCGCATCAGACACATTTGTTGCTTCCACTCTCATCGAATATGGTGG 2272
Db 2393 AGAAGAGAAATTTGGCGCATCTCTCTTTTGTGCGCATCAACACATTCCTTGAGAACGGAG 2452
QY 2273 TGTTCCTCAGTCTCCACATCCAGAAATCTCTTTTGAAGAAGCTPATCCATGCTATAAGTTG 2332
Db 2453 GACCCCTGAAGAGCGCAAGTCCAGCTTCTCTCTCTGAAGGAAGCTATACATGTTATCAGCTG 2512
QY 2333 TGCTATGAGGACAAAGTCTGAATGGGGAACAGATTTGGTTGGATCTATGGATCTGTAC 2392
Db 2513 CGGCTACGAGACAAAGCCGACTGGGGAAGAGATTGGCTGATTTACGATCGATCAC 2572
QY 2393 AGAAGATATCTTACTGGATTCAAGATGACGCAAGAGCTGGCGTTTCACTCTATTTCAT 2452
Db 2573 AGAGGATATCTTACTGGATTAAAGATGCACTGCCATGGCTGGCGTCTATTACTGCAAT 2632
QY 2453 GCCAAGCGCCAGCTTCAAGGATCTGCCCCCATCAATCTTTTCAATCGTCTGAACCA 2512
Db 2633 CCCAAGCGCCCTGCAATCAAGATTTCTGCGCTCTGAACCTTCCGACCGTCTTACCA 2692
QY 2513 AGTCTCGCGTGGGCTCTCGGTTCTGTGTAATCT--TTTCAGCGGCAATGCCCTTAT 2571
Db 2693 GGTCTCTCGTGGGCCCTTGGTCCGTCGAAATTTCTTTTCAGNAANCAATGNCACATTT 2752
QY 2572 GGTATGCTACGAGGCGCTCAAGTTCTGGAGATTCGTTATCATCAACACACCA 2631
Db 2753 GGTACGGATACGGCGGGCTAAATTTCTGGAAGGTTTCTTATCAATCCATCCATCG 2812
QY 2632 TTTTACCACATCACTCTCTCCGCTTCTAGTCTATTTATATGATTCCTGCTCTATCTGTCG 2691
Db 2813 TTTTATCCCTGAGCTCAATCTCTCTGCTTACTACTTCTGCTGCTGCTGCTGCTGCTG 2872
QY 2692 TCACCTGAAAGTTTCAATCATCCAGACATAGCAATTCGACATCTGTTTCAATTCGCG 2751
Db 2873 TCACGGGAAGTTTATCACACACAGCTTACCAATGTCGCGAGTATCTGTTTCAATGCGAC 2932
QY 2752 TCTTCCTTTCAATTTTCCGCACTCGTATCTTGGATGAGTGGAGTGGTGGGCAATG 2811
Db 2933 TTTTTCATCTGCATCTCCGTGACCGGCACTTCCGGAATGAGTGGAGTGGGTCGCAATCG 2992
QY 2812 ACAGTGTGGAGGAATCAACAGTCTTGGGTCTTGGAGTATCTTGCACATCTGTTTG 2871
Db 2993 ACAGTGTGGAGGAACAGCAGATTCGGGTCTATCGAGGGGTTTCGGGCACTGTTTCG 3052
QY 2872 CCGTCTTTTCAAGGTCTTCTGAGGTCTTGGCGGTATCGACCAATCTCATCTGTACCT 2931
Db 3053 CGGTGTTCCAGGGCTGCTGAAGGTGTTGGCGGCACTCGACACAGAGCTTCAACCGTGAGT 3112
QY 2932 CAAGGCTAATGACGAAGAGCGACTTGTGAGCTCTACATCTTCAAGTGGAGCGGC 2991
Db 3113 CGAAGGCGGGGACGACGA---GGAGTCTCGGAGCTGTACAGTTCAGTGGAGCCACC 3169
QY 2992 TTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGTTGGTGTCTGCTGCGCACT 3051
Db 3170 TGCTGATACCCCGACACAGCTCTCTCTGCTGAACTTCATCGGGTGTGGTGGCGGATCT 3229
QY 3052 CCTACGCAATCAACAGTGGTTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTCTTTG 3111
Db 3230 CGAACCGCATCAACAAACGGGTACGAGTCTGGGGCCCTCTTTCGGGAAGCTCTTCTTTG 3289
QY 3112 CTTTCTGGGTGATTGTTTCACTTATACCATCTTCAAGGTCTTATGGGAGGAAAC 3171
Db 3290 GGTTCGGGTGATCGTCAACCTGTATCCCGTCTCTCAAGGGTCTGTTGGGGAGGCAACA 3349
QY 3172 GCACACCGCATGTCATCTGGGTGTCCTCTCTGCTTCTATCTTCTCTTCTGCTGT 3231
Db 3350 GGACGCGACGATCGTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3409
QY 3232 GGGTTCGTTGATCCATTCATACCGTCTGCTGCGCCCAATATCCAAACCTGTGCA 3291
Db 3410 GGGTCCGCTGACACCGTCTCTCGCCAAAGAGCGACGGCCGCTCTCTGAGGAGTGTGCCC 3469
```

```
QY 3292 TCAACTGCTAGGAAGTGGAG 3313
Db 3470 TGGACTGCAACTGAAGTGGGG 3491

RESULT 7
AV104730 Zea mays PCO100501 mRNA sequence. 3783 bp mRNA linear HTC 16-OCT-2002
LOCUS Zea mays
DEFINITION Zea mays PCO100501 mRNA sequence.
ACCESSION AY104730
VERSION AY104730.1 GI:21207808
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3783)
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3783)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
1..3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

```
ORIGIN
Query Match 36.3%; Score 1315.6; DB 11; Length 3783;
Best Local Similarity 64.4%; Pred. No. 6.7e-264;
Matches 2132; Conservative 0; Mismatches 1074; Indels 102; Gaps 7;
QY 47 GCTCTGCGCGCCCATGACGCGGACGCGGCGCCCTGAAGTCCGGAGGACCGGGCCGG 106
Db 233 GCTCGTCTCATCCCGCGATGCGAGCCGAGCCGAAAGCCATGGACCGGAAACGG 292
QY 107 GGACGTGTGCCAGATCTGCGCGGACGCGCTGGGACACACGTTGGACGCGGAGCTTCAC 166
Db 293 CCAGGTGTGCCAGATTTTGGCGGACGAGCTGGGGCGCAACCCCGGAGGAGCCGTTCT 352
QY 167 GCGCTGGAGCTCTGCGGCTTCCCGGCTGCGCGCCCTGCTACGAGACGAGCGCAAGGA 226
Db 353 GGCCTGCAACAGTGTGGCTTCCCACTCTCGGGAGCTGCTACGAGTACGAGCGCGGA 412
QY 227 GGGCAACCCAGCGCTCTCCAGTGCAGACCAAGTGAAGCGCCACAGAGGAGCCAGC 286
Db 413 GGGCAGCAGAACTGCCCGCCAGTGCAGACCCGCTTCAAGGCCCTCAAGGGGTGCGCGG 472
QY 287 GATCCGGGGGAGGAGGACGACGACTGATGCCGATGATGTAGTACTTCAACTACC 346
```


QY	2714	AGAGATTAGCAACTTGGCCAGGTAATCTGGTTCAATTCGGCTCTCTTCCCTTTCAATTTTCGCCAC	2773
DB	2828	AGAGCTGAATAATGTTGGCCAGCCTGTGGTTCATGTCACATTTTATCTGCATTTTTCGCTAC	2887
QY	2774	TGGTATCCTTGACATGAGGTGGAGTGGTGTGGCAATTGACAGTGGTGGAGGAATGAACA	2833
DB	2888	GAGCATCCTAGAAATAGATGGAGTGGTGTGGAAATTGATGCTGGTGGAGGANTGAGCA	2947
QY	2834	GTTCCTGGGTCAATTTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTTCAGGGTCTTCTGA	2893
DB	2948	GTTCGGGTCAATTTGGAGGTGTCTCTCACACCTCTTTGCTGTCTTCAGGGACTTCTCAA	3007
QY	2894	GGTCTTGGCCGGTATCGACACCAACTTCATGTCACCTCAAGGCTCAATGACCAAGAGG	2953
DB	3008	GGTCATAGCTGGTGTATGATCAAGCTTTCACCGTGACATCAAGG---GTGGAGATGATGA	3064
QY	2954	CGACTTGTCTGAGCTCTACATGTTTCAAGTGGAGCAGCGCTTCTCATCCTCCGACGACCAT	3013
DB	3065	GGAGTCTCAGAGCTATATACATTCAATAGCATACCTTATTGATACCTCTACCACTT	3124
QY	3014	TTTGATCAATTAACATGGTGTGGTTCGTTGTCGCACTCTCTAGCCATCAACAGTGGTTA	3073
DB	3125	GCTTCTATTGAACCTTCATGCTGGTGGTCTGGCGTTTCAATTCGCGATCAATAACGATA	3184
QY	3074	CCAATCATGGGGCCGCTCTTTGGGAAGCTCTTCTTGGAGCTCTTCTGGGTGATTGTTCACTT	3133
DB	3185	TGAGTCATGGGGCCCTCTTTGGGAAGCTATCTTTTGGAGTGTCTTTGGGTGATTGTCATCT	3244
QY	3134	ATACCCATCTCTCAAGGCTCTTATGGGCGAGCAAAACCGCACCGATGTCATCGT	3193
DB	3245	TTATCCCTTTCTCAAGGTTGGTGGAGGCAAAACGAGACCAACGATTTGTCATCGT	3304
QY	3194	CTGGGCTGTCTCTCGCTTCTATCTCTCTCTGCTGTGGGTTCGTGTGATCCATTCAC	3253
DB	3305	CTGGTCCATTCTGCTGGCTTCAATCTCTCTGCTCTTTGGGTTCGGATTGATCCTTTTCT	3364
QY	3254	TACCGTCTCGTGGCCAAATATCCAAACCTGTGGCATCAACTGTAGGAAAGTGGAG	3313
DB	3365	TGGGAAGATGATGTCGCTCTTGGAGAGTGTGGTTGGATTGCAACTAGGATGTCAG	3424
QY	3314	T	3314
DB	3425	T	3425
RESULT 9			
LOCUS	CD726831	1874 bp	linear
DEFINITION	EST027 Cucurbita pepo testa subtracted cDNA Cucurbita pepo cDNA		EST 26-JUN-2003
ACCESSION	CD726831		clone CES similar to Cellulose synthase, mRNA sequence.
VERSION	CD726831.1	GI:32277678	
KEYWORDS	EST.		
SOURCE	Cucurbita pepo		
ORGANISM	Cucurbita pepo		
REFERENCE			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS			Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE			rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
JOURNAL			1 (bases 1 to 1874)
COMMENT			Bezold,T.N., Mathews,D., Loy,J.B. and Minocha,S.C. Molecular analysis of the hull-less seed trait in pumpkin: Expression profiles of cell wall related genes during development Unpublished (2003) Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold Dr. Minocha University of New Hampshire Rudman Hall, Durham, NH 03824, USA Tel: 603 862 3840 Fax: 603 862 3784 Email: sminocho@cisunix.unh.edu Degenerate primers and Taq were used to amplify cDNA for TOPO TA (Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three times using the Dyanamic ET Terminator Sequencing kit (Amersham

1796 CCTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAAGTTCCTCCCAAGAGTTTGA 1855
Db |||||
781 CTTGATGGACCCCAAGTGTGCCGAGATGATGCTATGTTCAAGTTTCTCAGAGTTTGA 840
QY |||||
1856 TGGGATTGATAGGATCATGATATGATAAGCAAGCAACACTGTCTTTTTGATATTAACCT 1915
Db |||||
841 TGGCAITGATCGAAGTATCGATATGCAATCGCAACACAGCTTTTCTTCGATGTGAACAT 900
QY |||||
1916 GAGGGGCTTACGGCAATCAAGGACCAAGTTTATGTGGGAACGTGTGCTTTTCAACAG 1975
Db |||||
901 GAAGGATTGGATGGCACTCAGGACCAAGTTTATGTTGGTACAGGTTGTGTTTCAATAG 960
QY |||||
1976 ACAGCTATCTGTTATGAGCCCCCAATTAAGGCGAAGAACCC--AGGTTTCTTGGC 2032
Db |||||
961 ACAGGCACCTTATGGCTACGACCTCTCTCCATGCTAGCTTATCTAAGACATCCCTCATC 2020
QY |||||
2033 ATCACTATGTGGGGGCAAGGAAGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAA 2092
Db |||||
1021 ATCTTGCTCTGTGTGGCTGTGCTCTTGTCTGTCTCTTCCCTTCCAAGAAAGATCTCAAAAGA 1080
QY |||||
2093 AAAGTCGAACAAGCATGTGACAGTTCTGTCTCAGTATTCAACTCGAAGACATAGAGGA 2152
Db |||||
1081 TCCGACTGAGATTCAGAGAGATCGAAAAGAGAGAGCTGTGCTGCAATCTTTAACT 1140
QY |||||
2153 GGGTGTGAAGTGTCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTT 2212
Db |||||
1141 CAGGGAATAGATAATATGATGAGTATGAGAGATCAATCTGATTTCTCACTGAGCTT 1200
QY |||||
2213 ACAGAGAGATTGGCCAGTCAGCAGCATTTGTTGCTCCACTCTCATGGAATATGCTGG 2272
Db |||||
1201 TGAGAAACTTTTCGGATTGCTGTCTGTGTTATCGAGTCTAGCTAATGGAATGGCGG 1260
QY |||||
2273 TGTTCCTCAGTCTCCACTCCAGAACTCTTTTGAAGAGCTATCCATGTCTATAAGTTG 2332
Db |||||
1261 AGTTGGGAATCTGAAATCCCTCGACTTGATCAAGGAAGCAATTCATGTCTATTAGCTG 1320
QY |||||
2333 TGGCTATGAGGCAAGTCTGAATGGGAACTGAGATTGCTGTGATCTATGGATCTGTCTAC 2392
Db |||||
1321 TGGTTATGAAGAGAAGTCCACTTGGGAAAAGAGATTGTTGGATATATGGGTGAGTGAC 1380
QY |||||
2393 AGAAGATATTCTTACTGGATTCAAGATGACGAAAGAGCTGGCGTTCACTATGTCAT 2452
Db |||||
1381 TGAGATATCTTAAACGGGTTTCAAGATGCAATTCGCGAGGGTGGAGTCCATCTACTGCAT 1440
QY |||||
2453 GCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCCATCAATCTTTCAGATCGTGAACCA 2512
Db |||||
1441 GCCATTTAGCCAGCATTCAGAGGTGCCCAACCAATTAACCTTTCTGATCGTCTCCACCA 1500
QY |||||
2513 AGTGTGCGGTGGGCTCTGGGTTCTGTTGAAATCTTTTCAGCCGGCAATGGCCCTTATG 2572
Db |||||
1501 AGTTCTTCGATGGGCACCTGGATCTGTTGAGATTTTCTTAGCACACACTGTCCATTAATG 1560
QY |||||
2573 GTATGGCT--ACGAGAGGGGCTCAAGTTCTGGAGAGATTCGCTTACATCAACACCAC 2629
Db |||||
1561 GTATGGAATTCGAGCGCGCGCTCAAAATGGCTCCAAAGAAATGGCTTACATAAAACACCAT 1620
QY |||||
2630 CATTTACCCCATTAACCTCTCTCCCGCTCTAGTCTATTGTTATATGCTGCTGCTATCTGCT 2689
Db |||||
1621 TGTCTATCCCTTCACTCGCTCCCTCTGTTGCTTACTGCTCATTCGCTGCAATCTGCT 1680
QY |||||
2690 GCTCATGGAAGTTTCACTATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCATATGC 2749
Db |||||
1681 GCTCACAGGAAAGTTTCACTATCCAAACGCTCTCGAACCTAGCAAGTACCCCTTTTCTCGG 1740
QY |||||
2750 GCTCTTCTTTCAATTTTCGGCATGTTGATCTTGGAGATGAGGTGGAGTGGTGGCAT 2809
Db |||||
1741 TCTCTTCTTGTCCATCAATCTCAAGAGTGTCTCGAGTGGTGGAGTGGTGGTATAGCAT 1800
QY |||||
2810 TGACAGTGGTGGAGGATGACAGTCTCTGGGTCAATGGAGGTATCTCTGACATCTGTT 2869
Db |||||
1801 CGAAGATATATGGGTACAGAGCAATCTGGGTAAATCGGAGGCGTCCCGGCACATCTCTT 1860
QY |||||
2870 TGCGGTCTTTCCAGG 2883

Db |||||
1861 TGCGGTCTTCCAAG 1874

RESULT 10

BM816138

LOCUS

DEFINITION

BM816138 880 bp mRNA linear EST 05-MAR-2002
HC109B12_T3.ab1 HC Hordeum vulgare subsp. vulgare cDNA clone
HC109B12_T3.ab1 similar to (AF200528) cellulose synthase-4 [Zea
mays], (AF200529) cellulose synthase-5 [Zea mays], (AF200533)
cellulose synthase-9 [Zea mays], cellulose synthase catalytic
subunit [Arabidopsis thaliana], unnamed protein product [Arabidopsis
thaliana], mRNA sequence.

ACCESSION

BM816138

VERSION

BM816138.1

KEYWORDS

EST.

SOURCE

Hordeum vulgare subsp. vulgare

ORGANISM

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Hordeum.

1 (bases 1 to 880)

Ozturk,N.Z., Michalowski,C.B., Brazille,S., Borchert,C.,

Palacio,C., Normand,C., Murphy,C., Kelley,R., Sant,S.A.,

McLaughlin,H., Fredricksen,M.A. and Bohnert,H.J.

Monitoring large-scale changes in transcript abundance in drought-

and salt-stressed barley

Unpublished (2002)

Contact: Mark A. Fredricksen

Plant Biology

University of Illinois

1201 W Gregory Dr, Urbana, IL 61801, USA

Tel: 2172655473

Email: bohnertlab@life.uiuc.edu.

FEATURES

Location/Qualifiers

1..880

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/strain="cv tokak"

/sub_species="vulgare"

/db_xref="taxon:112509"

/clone="HC109B12_T3.ab1"

/tissue_type="Root"

/dev_stage="3 week old"

/clone_lib="HC"

/note="6 and 10 hour drought stress by placing plants on

moist paper (75% rel. humidity) in light"

ORIGIN

Query Match

Best Local Similarity

Matches

801; Conservative

0; Mismatches

30; Indels

2; Gaps

2;

QY 2469 TTCAAGGATCTGCCCCCATCAATCTTCAGATCGTCTGAACCAAGTCTGCGTGGGCT 2528

Db |||||

48 TTCAAGGATCTGCCCCCATCAATCTTCAGATCGTCTGAACCAAGTCTGCGTGGGCT 106

QY 2529 CTCGGTCTGTTGAAATCTTTTCAGCGGCAATTCGCCCTTATGGTATGGCTACGGAGG 2588

Db |||||

107 CTTCGGTCCGTTGAATCTTTTCAGCGGCAATTCGCCCTTATGGTATGGCTACGGAGG 166

QY 2589 CGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCAACCATTTACCCATAACCTCT 2648

Db |||||

167 CGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCAACCATTTACCCATAACCTCT 226

QY 2649 CTCGGCTTCTAGTCTATGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2708

Db |||||

227 CTCGGCTTCTAGTCTATGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 286

QY 2709 ATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCATTTGGCTCTTCTTCAATTTTC 2768

Db |||||

287 ATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCATTTGGCTCTTCTTCAATTTTC 346


```

RESULT 12
BQ802778      762 bp      mRNA      linear      EST 30-JUL-2002
LOCUS        WHE2829_H06_O11ZS Triticum monococcum vernalized apex cDNA library
DEFINITION   Triticum monococcum cDNA clone WHE2829_H06_O11, mRNA sequence.
ACCESSION   BQ802778
VERSION     BQ802778.1 GI:22017747
KEYWORDS     EST.
SOURCE      Triticum monococcum
ORGANISM    Triticum monococcum
REFERENCE   1 (bases 1 to 762)
AUTHORS     Anderson,O.D., Chao,S., Crossman,C., Dubcovsky,J., Echenique,V.,
            Lazo,G.R., Pham,J., Rausch,C.J., Stamova,B., Wilson,C. and Woo,J.
TITLE       The structure and function of the expressed portion of the wheat
            genomes - Vernalized apex cDNA library from Triticum monococcum
            Unpublished (2002)
JOURNAL
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
            Location/Qualifiers
              1..762
                /organism="Triticum monococcum"
                /mol_type="mRNA"
                /cultivar="G3116"
                /db_xref="taxon:4568"
                /clone="WHE2829_H06_O11"
                /tissue_type="Vernalized apex"
                /dev_stage="One month old plants"
                /lab_host="E. coli XL0LR"
                /clone_lib="Triticum monococcum vernalized apex cDNA
                library"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                Site 1: EcoRI; Site 2: XhoI; One-month old plants were
                subjected to vernalization treatment by placing them in
                the cold room at 6 C, under 15hr light/9hr dark condition.
                Total RNA was prepared from apex tissue extracted from
                plants with no cold treatment; and from plants with
                2-week, 4-week and 6-week cold treatment separately. Equal
                amount of total RNA was pooled from all four samples, a
                cDNA library was made using pooled polyA RNA and cDNA
                clones were in vivo excised at the University of
                California, Davis (V. Echenique, B. Stamova, J.
                Dubcovsky). Plasmid DNA preparations and DNA sequencing
                were performed in the OD Anderson lab (all other
                authors)."
```

Query Match 20.7%; Score 750.4; DB 13; Length 762;
 Best Local Similarity 99.9%; Pred. No. 6e-146;
 Matches 751; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 625 CACCAATCCGTCGAAGGAGTTCTCCGGCAGTATTGGGAATGTCCTGGAAGAGAGAG 684
Db 11 CAGCAATCCGTCGAAGGAGTTCTCCGGCAGTATTGGGAATGTCCTGGAAGAGAGAG 70
Qy 685 TTGATGGCTGGAATGAAGCAGACACAGGGTGGGATTCCTCCATGATTAATGGCAAGCA 744
Db 71 TTGATGGCTGGAATGAAGCAGACACAGGGTGGGATTCCTCCATGATTAATGGCAAGCA 130
Qy 745 TTGCTCCCTCTGAAGGTCGGCAGCTACTGACATGATGATCTACTGAATACACATGG 804
Db 131 TTGCTCCCTCTGAAGGTCGGCAGCTACTGACATGATGATCTACTGAATACACATGG 190
```

ORIGIN

```

805 AAGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTACTAGAAAGTCCCATGCTT 864
191 AAGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTACTAGAAAGTCCCATGCTT 250
865 CTCCAAAATAAATCCCTACAGAAATGGTCAATGTTCTGCGGTGGTGTCTTAAGCATCT 924
251 CTCCAAAATAAATCCCTACAGAAATGGTCAATGTTCTGCGGTGGTGTCTTAAGCATCT 310
925 TCTGCACTACCGTCTCACAATAATCCCTGCTGCTAAATGCATACACCACTGTGGCTTTTATCTG 984
311 TCTGCACTACCGTCTCACAATAATCCCTGCTGCTAAATGCATACACCACTGTGGCTTTTATCTG 370
985 TTATATGTGAGATTTGGTTTGGCTTTTAACTGATATCTGATATCGATCAGTTCCTCCGAAGTGGTTTC 1044
371 TTATATGTGAGATTTGGTTTGGCTTTTAACTGATATCTGATATCGATCAGTTCCTCCGAAGTGGTTTC 430
1045 CAATCAACGGGAGACCTACCTTGATAGACTGGCTTTTAAGGTATGACCGAGAGGTGAAC 1104
431 CAATCAACGGGAGACCTACCTTGATAGACTGGCTTTTAAGGTATGACCGAGAGGTGAAC 490
1105 CGTCTCAGTTGGCTGCTGTTGACATATTTGTGACATACAGTTCGACACCTTTGAAGAGGCCAC 1164
491 CGTCTCAGTTGGCTGCTGTTGACATATTTGTGACATACAGTTCGACACCTTTGAAGAGGCCAC 550
1165 CTATCGTCACTCCCAACACTGTGCTATCCATTCTTGGCTGTGATTCATTCCTGGGACAGG 1224
551 CTATCGTCACTCCCAACACTGTGCTATCCATTCTTGGCTGTGATTCATTCCTGGGACAGG 610
1225 TCTCTTGTATGATGATGATGACGAGCTTCAATGCTGATTTTGACGATTTGGCTGAGA 1284
611 TCTCTTGTATGATGATGATGACGAGCTTCAATGCTGATTTTGACGATTTGGCTGAGA 670
1285 CTTACAGATTTGCTAGGAAATGGGTACCATTTGTGAAGAAGTATGACATTGAACCCAGAG 1344
671 CTTACAGATTTGCTAGGAAATGGGTACCATTTGTGAAGAAGTATGACATTGAACCCAGAG 730
1345 CTCGCGAGTTTACCTTTGCCAGAAAAATGTAT 1376
731 CTCGCGAGTTTACCTTTGCCAGAAAAATGTAT 762
```

RESULT 13
 BQ368813 870 bp mRNA linear EST 22-OCT-2001
 LOCUS HVSME10020M13f Hordeum vulgare 20 DAP spike EST library HVCNDA0010
 DEFINITION (20 DAP) Hordeum vulgare subsp. vulgare cDNA clone HVSME10020M13f,
 mRNA sequence.
 ACCESSION BQ368813 GI:13257914
 VERSION BQ368813.1
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 870)
 AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D.,
 Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W.,
 Fenton, R.D., Close, S.J., Oates, R. and Main, D.
 TITLE Development of a genetically and physically anchored EST resource
 JOURNAL for barley genomics: Morex 20 DAP spike cDNA library
 COMMENT Unpublished (2001)
 Contact: Wing RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 537
 Seq primer: AATTAACTCCTCACTAAAGG
 High quality sequence start: 777.

FEATURES

source

Location/Qualifiers
 1..870
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /sub_species="vulgare"
 /db_xref="taxon:112503"
 /clone="HVSMEI0020M131"
 /tissue_type="20 DAP spike"
 /lab_host="SOLR"
 /clone_lib="Hordeum vulgare 20 DAP spike EST library
 HVCDA0010 (20 DAP)"
 /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
 Plants were grown in the greenhouse at the University of
 California, Riverside (Fenton, SJ Close, TJ Close). Whole
 spikes with awns trimmed were collected at 20 DAP
 (Fenton). Total RNA was prepared, poly(A) RNA was
 purified, one primary unamplified cDNA library was made,
 and 1 million pfu were in vivo excised to give pluescript
 SK(-) cDNA phagemids in the TJ Close lab at the University
 of California, Riverside (Choi). Phagemids were plated and
 picked at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
 Oates, Rambo, Main). The sequence has been trimmed to
 remove vector sequence and contains a minimum of 100 bases
 of phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders
 see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

Query Match 20.1%; Score 729.4; DB 12; Length 870;
 Best Local Similarity 93.8%; Pred. No. 1.5e-141; Indels 6; Gaps 5;
 Matches 813; Conservative 0; Mismatches 48

QY 1747 GTGATCACTACATCAACACAGCAGGCTGTCCGAGAGCTATGTCTCTTAATGGATC 1806
 Db 1 GTGATCACTACATCAACACAGCAGGCTGTCCGAGAGCTATGTCTCTTAATGGATC 60

QY 1807 CAAACCTAGTCCCAAGTCTGTATGTGCAGTTCACCAAGGTTTGATGGGATTCATA 1866
 Db 61 CAAACCTAGTCCCAAGTCTGTATGTGCAGTTCACCAAGGTTTGATGGGATTCATA 120

QY 1867 GGAATGATGATATGCAACAGGAACACTGTCTTTTGTATTAACCTGAGGGGCTTG 1926
 Db 121 GGAATGATGATATGCAACAGGAACACTGTCTTTTGTATTAACCTGAGGGGCTTG 180

QY 1927 ACGGCATTCAAGGACCACTTTATGTGGAACTGTGTGTTTTCACAGAACAGCTATCT 1986
 Db 181 ATGGCATTCAAGGACCACTTTATGTGGAACTGTGTGTTTTCACAGAACAGCTATCT 240

QY 1987 ATGTTTATGAGCCCCCAATTAAAGCGGAAGACCAAGTTCCTGGCATCACTATGTGGG 2046
 Db 241 ATGTTTATGAGCCCCCAATTAAAGCGGAAGACCAAGTTCCTGGCATCACTATGTGGG 300

QY 2047 GCAGAGAGAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAGTCGAACAAGC 2106
 Db 301 GCAAGAGAGAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAGTCAACAAGC 360

QY 2107 ATGTGGACAGTCTGTCTCCAGTATTCAATCTCGAAGACATAGAGGAGGTGTGAAGTG 2166
 Db 361 ATGTGGACAGTCTGTCTCCAGTATTCAATCTCGAAGACATAGAGGAGGTGTGAAGTG 420

QY 2167 CTGGTTTGTATGATGAGAAATCAGTTCTCATGTCTCAATAGCTTAGAGAGAGATTG 2226
 Db 421 CTGGTTTGTATGATGAGAAATCAGTTCTCATGTCTCAATAGCTTAGAGAGAGATTG 480

QY 2227 GCCAGTCAGCAGCAATTTGTTGCTCCACTCTCATGGAATATGGTGTGTTCTCAGTCT 2286
 Db 481 GCCAGTCAGCAGCAATTTGTTGCTCCACTCTCATGGAATATGGTGTGTTCTCAGTCT 540

QY 2287 CCATCTCCAGAACTCTTTTGAAGAGCTATCCATCTCATAGTTCGTGCTATGAGGACA 2346
 Db 541 CCATCTCCAGAACTCTTTTGAAGAGCTATCCATCTCATAGTTCGTGCTATGAGGACA 599

QY 2347 AGTCTGAATGGGGAAGTGAATTTGGTATCTATGGATCTCTGCACAGAGATATTCTTA 2406
 Db 600 AGTCTGAATGGGGAAGTGAATTTGGTATCTATGGATCTCTGCACAGAGATATTCTTA 659

QY 2407 CTGGATTCAAGATGACGACGAGAGGCTGGC-GTTAGTCTATTGATGTCGCCAGCGCCA 2465
 Db 660 CTGGATTCAAGATGACGACGAGAGGCTGGC-GTTAGTCTATTGATGTCGCCAGCGCCA 719

QY 2466 GCTTTCAAGGATCTGCCCTCATCATCTTTCAGATCTGTCGACCAAGTCTGCGGTGG 2525
 Db 720 GCTTTCAAGGATCTGCCCTCATCATCTTTCAGATCTGTCGACCAAGTCTGCGGTGG 779

QY 2526 GCTCTCGGTTCTGTTCAAAATCTTTTCAGCGGCAATGGCCCTTATGGTATGGCTAGGA 2585
 Db 780 GCTCTCGGTTCTGTTCAAAATCTTTTCAGCGGCAATGGCCCTTATGGTATGGCTAGGA 836

QY 2586 GGGCGGCTCAAGTCTCGGAGATTC 2612
 Db 837 AGGCCCT-AAGTCTCTGGAAAAATCC 862

RESULT 14

CD878620/c

LOCUS

DEFINITION

CD878620

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..723

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="recital"

/db_xref="taxon:4565"

/clone="AZO4103D19"

/tissue_type="root"

/clone_lib="AZO4"

ORIGIN

Query Match 19.2%; Score 697.8; DB 14; Length 723;
 Best Local Similarity 93.6%; Pred. No. 5.8e-135; Indels 1; Gaps 1;
 Matches 710; Conservative 0; Mismatches 2

QY 2914 CCAACTTCACCTGTCACCTCAAGAGCTATGACGAAGAGGCGACTTTGCTGAGCTTACA 2973
 Db 712 CCAACTTCACCTGTCACCTCAAGAGCTATGACGAAGAGGCGACTTTGCTGAGCTTACA 654

```
QY 2974 TGTTCAGTGGACGACGCTTCTCATCCCTCCGACGACCAATTTTGATCAATTAACATGGTTG 3033
Db 653 TGTTCAGTGGACGACGCTTCTCATCCCTCCGACGACCAATTTTGATCAATTAACATGGTTG 594
QY 3034 GTGTCTGCTTGGCAGCTCTCTAGCCATCAAGAGTGGTTACCAATCATGGGGCCGCTCT 3093
Db 593 GTGTCTGCTTGGCAGCTCTCTAGCCATCAAGAGTGGTTACCAATCATGGGGCCGCTCT 534
QY 3094 TTGGGAAGCTCTTCTTTGGCTTCTGGGTGATTTCTCACTATATACCAATCCCTCAAGGGTC 3153
Db 533 TTGGGAAGCTCTTCTTTGGCTTCTGGGTGATTTCTCACTATATACCAATCCCTCAAGGGTC 474
QY 3154 TTATGGGAGGCAAAACCCACACACGAGATTGTCTGCTGGGTGCTCTCTCGCTT 3213
Db 473 TTATGGGAGGCAAAACCCACACACGAGATTGTCTGCTGGGTGCTCTCTCGCTT 414
QY 3214 CTATCTTCTCTTGGTGGTTCGTGTGATTCATTCATCCCTCGTGGGCCAA 3273
Db 413 CTATCTTCTCTTGGTGGTTCGTGTGATTCATTCATCCCTCGTGGGCCAA 354
QY 3274 ATATCCAAACCTGTGGCATACTGCTAGGAAAGTGGAGTTTGTAGACAGAAATAT 3333
Db 353 ATATCCAAACCTGTGGCATACTGCTAGGAAAGTGGAGTTTGTAGACAGAAATAT 294
QY 3334 AACAGTATCGAGCAACAAACCCGCGAGCCAGAGATATTTATGTTGGGTTGTGAATTA 3393
Db 293 AACAGTATCGAGCAACAAACCCGCGAGCCAGAGATATTTATGTTGGGTTGTGAATTA 234
QY 3394 CTACGTTTGGAGAACTTGTCAAATTTGAGAAACACATTTGTAATAGATGATAGACC 3453
Db 233 CTACGTTTGGAGAACTTGTCAAATTTGAGAAACACATTTGTAATAGATGATAGACC 174
QY 3454 ATCTACCGTTTTCATGAGTAAAGCTCTCTTTTGGAAACAAAGGAATCTCATTTGTA 3513
Db 173 ATCTACCGTTTTCATGAGTAAAGCTCTCTTTTGGAAACAAAGGAATCTCATTTGTA 114
QY 3514 AACCTATAGGAATTTTCTATGAGCACTTTGGATTGTAGGAATGACCTATGAATGTT 3573
Db 113 AACCTATAGGAATTTTCTATGAGCACTTTGGATTGTAGGAATGACCTATGAATGTT 54
QY 3574 GTATTTATTTATATATAAATTTCCCTGCTTCCACATTTTGGAGGATTTT 3626
Db 53 GTATTTATTTATATATAAATTTCCCTGCTTCCACATTTTGGAGGATTTT 1

RESULT 15
LOCUS CD930401
DEFINITION GR45.111C19F010418 GR45 Triticum aestivum cDNA clone GR45111C19,
mRNA sequence.
ACCESSION CD930401
VERSION CD930401.1 GI:32778165
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 701)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
Location/Qualifiers
1..701
/organism="Triticum aestivum"
```

```
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="GR45111C19"
/tissue_type="grain (45 degrees per day after
pollination)"
/clone_lib="GR45"
```

ORIGIN

```
Query Match
Best Local Similarity 99.0%; Score 689.8; DB 14; Length 701;
Matches 694; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2324 CATAGTTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGATCTATGG 2383
Db 1 CATAGTTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGATCTATGG 60
QY 2384 ATCTGTCAAGAGATATTTCTTACTGGATTCAAGATGCAAGGAGGCTGGGTTCACT 2443
Db 61 ATCTGTCAAGAGATATTTCTTACTGGATTCAAGATGCAAGGAGGCTGGGTTCACT 120
QY 2444 CTATTGCATGCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCATCAATCTTTTCAGATCG 2503
Db 121 CTATTGCATGCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCATCAATCTTTTCAGATCG 180
QY 2504 TCTGAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTGAAATTTCTTTTCAGCCGGCATTG 2563
Db 181 TCTGAATCAAGTGTGCGGTGGGCTCTCGGTTCTGTGAAATTTCTTTTCAGCCGGCATTG 240
QY 2564 CCCCTTATGGTATGCTACGAGGGGGCTCAAGTTCCTGGAGAGATTCGCTTACATCAA 2623
Db 241 CCCCTTATGGTATGCTACGAGGGGGCTCAAGTTCCTGGAGAGATTCGCTTACATCAA 300
QY 2624 CACCAACATTTACCACATAACCTCTCTCCCGCTTCTAGTCTATTTATATTCCTGCTAT 2683
Db 301 CACCAACATTTACCACATAACCTCTCTCCCGCTTCTAGTCTATTTATATTCCTGCTAT 360
QY 2684 CTGTCTGTCTACTGAAAGTTTCATGCCAGAGATTAGCAACTTTGGCCAGTATCTGGTT 2743
Db 361 CTGTCTGTCTACTGAAAGTTTCATGCCAGAGATTAGCAACTTTGGCCAGTATCTGGTT 420
QY 2744 CATTGCGCTCTTCTTTTCAATTTTCCACCTGATCTTCTGAGATGAGGTGGTGT 2803
Db 421 CATTGCGCTCTTCTTTTCAATTTTCCACCTGATCTTCTGAGATGAGGTGGTGT 480
QY 2804 TGGCATTGACGAGTGGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACA 2863
Db 481 TGGCATTGACGAGTGGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACA 540
QY 2864 TCTGTTTGGCGCTCTTTTCAAGGTCTTCTGAAAGTCTTGGCGGTATCGACACCACTTCAC 2923
Db 541 TCTGTTTGGCGCTCTTTTCAAGGTCTTCTGAAAGTCTTGGCGGTATCGACACCACTTCAC 600
QY 2924 TGTCACTCAAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTTACATGTTCAAGTG 2983
Db 601 TGTCACTCAAAGGCTAATGATGAAGGCGCACTTTGCTGAGCTTACATGTTTCAAGTG 660
QY 2984 GACGACGCTTCTCATCCCTCCGACGACCAATTTTGATCATTA 3024
Db 661 GACGACGCTTCTCATCCCTCCGACGACCAATTTTGATCATTA 701
```

Search completed: August 23, 2004, 07:56:30
Job time : 8691 secs

1	4499	77.9	1065	4	US-09-221-013A-10	Sequence 10, Appl
2	4455.5	77.1	881	3	US-08-960-048-8	Sequence 8, Appl
3	4455.5	77.1	881	4	US-09-838-586-8	Sequence 8, Appl
4	3963.5	68.6	1081	4	US-09-221-013A-6	Sequence 6, Appl
5	3959.5	68.5	1081	4	US-09-221-013A-12	Sequence 12, Appl
6	3633.5	62.9	1084	4	US-09-221-013A-8	Sequence 8, Appl
7	3477	60.2	974	4	US-08-960-048-6	Sequence 6, Appl
8	3477	60.2	974	3	US-08-938-586-6	Sequence 6, Appl
9	2811	48.7	685	3	US-08-960-048-7	Sequence 7, Appl
10	2811	48.7	685	4	US-09-838-586-7	Sequence 7, Appl
11	2629.5	45.5	629	4	US-09-221-013A-2	Sequence 2, Appl
12	1756.5	30.4	547	4	US-09-221-013A-4	Sequence 4, Appl
13	282	4.9	57	4	US-08-812-008-48	Sequence 14, Appl
14	268	4.6	693	3	US-08-960-048-11	Sequence 11, Appl
15	268	4.6	693	4	US-09-838-586-11	Sequence 11, Appl
16	255.5	4.4	861	3	US-08-960-048-12	Sequence 12, Appl
17	255.5	4.4	861	4	US-09-838-586-12	Sequence 12, Appl
18	251.5	4.4	946	4	US-09-489-039A-11056	Sequence 11056, A
19	238.5	4.1	3031	1	US-07-689-008-2	Sequence 2, Appl
20	237.5	4.1	756	3	US-08-960-048-10	Sequence 2, Appl
21	237.5	4.1	756	4	US-09-147-236-3	Sequence 3, Appl
22	237.5	4.1	756	4	US-09-522-474-3	Sequence 3, Appl
23	237.5	4.1	756	4	US-09-838-586-10	Sequence 10, Appl
24	236.5	4.1	738	4	US-09-543-681A-7528	Sequence 7528, Ap
25	230	4.0	883	4	US-09-543-681A-6947	Sequence 6947, Ap
26	210.5	3.6	703	4	US-09-489-039A-10820	Sequence 10820, A
27	207	3.6	723	3	US-08-960-048-9	Sequence 9, Appl

Db	299	ALSWILDQFPVFNRETYDLRALRYDREGEPSQLAAVDIFVSTVDLPKPEPLVTANT	358
Qy	376	VLSILAVDYVPDKVSCVSDGDGASMTFDALARTSEFARKWPFVKYDIERAPREFYFC	435
Db	359	VLSILAVDYVPDKVSCVSDGDGAAMLSFSLAETSEFARKWPFVKYSEIAPREWPEFA	418
Qy	436	OKDYLKDKVQPSFKDRAAMKEYEFEKIRINALVSKALVPEEGWIMQDGTWPGNNT	495
Db	419	AKDYLKDKVQTSFVKDRAAMKEYEFEKIRINALVSKALCPEEGWIMQDGTWPGNNT	478
Qy	496	RDPGMIQVFLHSGGLDTEGNLPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNG	555
Db	479	GDHPGMIQVFLGQGGGLDAEGNLPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNG	538
Qy	556	QYMLNLDCDHYINNSKAVREAMCFMDPNLGPQVYVQPPQRPDGDINDRDRYANRNVFF	615
Db	539	PFILNLDCDHYINNSKALREAMCFMDPNLGKQVYVQPPQRPDGDINDRDRYANRNVFF	598
Qy	616	DINLRGLDGTQGPVYVGTGVFNRTALYGEPEPLKA--KKPGFLASI-CGGKKKASKSKKR	673
Db	599	DINLRGLDGIQGPVYVGTGVFNRTALYGEPEPLKVKKHKPSLLKLCGGSRKKKNSKAKK	658
Qy	674	SSDKKSKNHVDSVPVFNLEIDIEEGVEGAGFDDKSVLMSQMSLEKRFQCSAAFAVASTL	733
Db	659	ESDKKSGRHSTSTVPVFNLDIDIEEGVEGAGFDDKAKLLMSQMSLEKRFQCSAVFAVASTL	718
Qy	734	MEYGVQPOSTPESILKEAHHVISCGYEDKSEWGETIGWYGSVTEIDLTGFKMHARGWR	793
Db	719	MENGVPPSATPENFLKEAHHVISCGYEDKSDMGWEIGWYGSVTEIDLTGFKMHARGWR	778
Qy	794	SVYCMRKPEPAFKGSAPINLSRLNQVLRWALGSVEILFSRHCPLWYGYGRLKFLERFAY	853
Db	779	SIYCMKPLPAFKGSAPINLSRLNQVLRWALGSVEILFSRHCPITWYGYGRLKFLERFAY	838
Qy	854	INTTIPTUTSPLLYCYCILPAICLLTGKFMPEISNLASITWIALFUSIATGILEMRWS	913
Db	839	VNNTIPTYITSPLLMYCTLLAVCLTFNQFIIPQISNIAISWFLSFLUSIATGILEMRWS	898
Qy	914	GVGIDEWARNQFVWIGGISAHLFAVFGLLKVLAGIDTNFTVTSKANDEGOFAEIYMF	973
Db	899	GVGIDEWARNQFVWIGGVSAAHLFAVFGILKVLAGIDTNFTVTSKASDEGOFAEIYLF	958
Qy	974	KWTTLLIPTTILINMVGWVAGTSYAINSGYQSWGPLFGKLPFAFWVIVHLYPFLKGLM	1033
Db	959	KWTTLLIPTTILLIVNLGVWAGSYAINSGYQSWGPLFGKLPFAFWVIVHLYPFLKGLM	1018
Qy	1034	GRQNRPTPIVWAVLLASISFLLMVRVDPPTTRLAGNIOTCGINC	1080
Db	1019	GRONTPPIVWVSVLLASISFLLMVRIDPPTSRVTGDIILECGINC	1065

Query Match	77.1%;	Score	4455.5;	DB 3;	Length	881;			
Best Local Similarity	93.4%;	Pred. No.	0;						
Matches	823;	Conservative	39;	Mismatches	18;	Indels	1;	Gaps	17;

QY	201	GNVAKERV	DGWMKQDKGAI	PMTNGT	STISAP	SEGRAAT	ID	ASTEY	NMD	ALL	DET	RQP	260	
DB	1	GNVAKERV	DGWLKQDKGAI	PMTNGT	SIAP	SEGRGVGD	ID	AST	DYNN	D	ALL	DET	RQP	60
QY	261	LSRKVP	IASSKINP	RMVIL	BLVLS	IFLHV	RLTP	VRN	AYPL	WLLSV	ICET	WEAL	SWI	320
DB	61	LSRKVP	LPSSRINP	RVNIV	LVLVLS	IFLHV	RLTP	VRN	AYPL	WLLSV	ICET	WEAL	SWL	120
QY	321	LQDPFK	WPINRET	LDL	ALRY	DREG	EPSQL	AAVD	IFV	STVD	PLKE	PPIT	VANT	380
DB	121	LQDPFK	WPINRET	LDL	ALRY	DREG	EPSQL	AAVD	IFV	STVD	PMKE	PPIT	VANT	180
QY	381	AVDYP	VDKVCV	SDDG	ASML	TFDA	LAET	SEF	ARKV	VPFV	KKYD	IEPR	AP	440
DB	181	AVDYP	VDKVCV	SDDG	AAML	TFDA	LAET	SEF	ARKV	VPFV	KKYNI	IEPR	AP	240
QY	441	LKDKVP	PSVKD	RRAMK	REYEF	KIR	INAL	VS	KAL	KVP	EEG	WIMQ	DT	500
DB	241	LKDKVH	PSVKD	RRAMK	REYEF	KIR	ING	LA	VAK	KVP	EEG	WIMQ	DT	300
QY	501	MTQV	FLGHS	GGLD	TEG	NELP	RLVY	VS	REK	RPQ	QHHK	KAG	AMN	560
DB	301	MTQV	FLGHS	GGLD	TEG	NELP	RLVY	VS	REK	RPQ	QHHK	KAG	AMN	360
QY	561	LCD	HHYNN	S	KAV	RE	AMC	FL	MD	NL	PG	VCY	QV	620
DB	361	LCD	HHYNN	S	KAL	RE	AMC	FL	MD	NL	PG	VCY	QV	420
QY	621	GLD	GI	QPV	V	GTG	CV	N	RTA	I	Y	GE	Y	679
DB	421	GLD	GI	QPV	V	GTG	CV	N	RTA	I	Y	GE	Y	480
QY	680	SNKH	VD	SS	VPV	F	N	LED	IE	G	V	E	G	739
DB	481	SNKH	VD	SA	VPV	F	N	LED	IE	G	V	E	G	540
QY	740	PQS	TS	PE	SL	KE	AH	VI	SC	GE	Y	DK	TE	799
DB	541	PQS	TS	PE	SL	KE	AH	VI	SC	GE	Y	DK	TE	600
QY	800	KR	PA	FG	S	A	P	I	N	S	D	R	L	859
DB	601	KR	PA	FG	S	A	P	I	N	S	D	R	L	660
QY	860	PL	T	S	L	P	L	L	Y	C	I	L	P	919
DB	661	PL	T	S	L	P	L	L	Y	C	I	L	P	720
QY	920	W	N	E	Q	P	W	I	G	G	I	S	A	979
DB	721	W	N	E	Q	P	W	I	G	G	I	S	A	780
QY	980	I	P	T	T	I	L	I	N	M	V	G	V	103
DB	781	I	P	T	T	I	L	I	N	M	V	G	V	840
QY	1040	P	T	V	I	W	A	V	I	L	A	S	F	
DB	841	P	T	V	I	W	A	V	I	L	A	S	F	

RESULT 3
US-09-838-586-8
; Sequence 8, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences

```
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Oryzae sativa
US-09-838-586-8

Query Match      77.1%; Score 4455.5; DB 4; Length 881;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;

QY 201 GNVAKERVDMKMDKQGAIPMTNGTSTAPSEGRAATDIDASTEYNMEDALLNDETROP 260
DB 1 GNVAKERVDMKMDKQGAIPMTNGTSTAPSEGRGVGDIADSTYNNEDALLNDETROP 60
QY 261 LSRKVPASCKNPYRMVILRLVLSIFLHYRLTNPNVNAVPLMLLSVICEIFALSWI 320
DB 61 LSRKVPESRPINPYRNIVLRLVLSIFLHYRIITPNVNAVPLMLLSVICEIFALSWL 120
QY 321 LDQFPKWPPIRETLYDLRALAYDREGPSLAADVDFYSTVDPLKEPPIVANTVLSIL 380
DB 121 LDQFPKWPPIRETLYDLRALAYDREGPSLAADVDFYSTVDPMKEPPIVANTVLSIL 180
QY 381 AVDPYVDKVCYSVDDGASMLTFDALAETSEFARKWVPVKYDIEPRAPEYFQCKIDY 440
DB 181 AVDPYVDKVCYSVDDGASMLTFDALAETSEFARKWVPVKYDIEPRAPEYFQCKIDY 240
QY 441 LKDKVPQSFVKDRRAMKREYEFKIRINALVSKALKVPEEGIMQDGTWPWGNTRDHPG 500
DB 241 LKDKVHPSFVKDRRAMKREYEFKIRINGLVAQKVPPEEGIMQDGTWPWGNTRDHPG 300
QY 501 MIQVFLHSGGLDTSGNELPLRVYVSREKPGFQHHKAGANALVRVSAVLNQGMYLN 560
DB 301 MIQVFLHSGGLDTSGNELPLRVYVSREKPGFQHHKAGANALVRVSAVLNQGMYLN 360
QY 561 LDCDHYINNSKAVREAMCFMDPNLGPVYQFQRFQDGDIDNRYANRNTVFEDINLR 620
DB 361 LDCDHYINNSKALREAMCFMDPNLGRSVYQFQRFQDGDIDNRYANRNTVFEDINLR 420
QY 621 GLDGIQGPVYGTGCVFNRTALYGEYPTIKQKKGSLSSLCGRKKASKSKKSSDKKK 679
DB 421 GLDGIQGPVYGTGCVFNRTALYGEYPTIKQKKGSLSSLCGRKKASKSKKSSDKKK 480
QY 680 SNKHVDSVPVFNLEDEEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFPVASTLMEYGV 739
DB 481 SNKHVDSVPVFNLEDEEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFPVASTLMEYGV 540
QY 740 PQSSTPESLLKEAHIHIVISCGYEDKSEWTEIGWIYGSVTEIDITGFKMHARGWSVYCMP 799
DB 541 PQSATPESLLKEAHIHIVISCGYEDKTEWTEIGWIYGSVTEIDITGFKMHARGWSVYCMP 600
QY 800 KRPAFKGAPINLSDRLNQVLWALGSVEILFSRHCPWYGYGRLKFLERFAYINTTIY 859
DB 601 KRPAFKGAPINLSDRLNQVLWALGSVEILFSRHCPWYGYGRLKFLERFAYINTTIY 660
QY 860 PLTSLPLAVYCTLPAICLLTGKFMPEJSNLASIMFIALFSIFATGILEMRWSVGIDE 919
DB 661 PLTSLPLAVYCTLPAICLLTGKFMPEJSNLASIMFIALFSIFATGILEMRWSVGIDE 720
QY 920 WWRNEQFVIGISAHLPFAVQGLKVLAGIDTNTFTVTSKANDERGDFAEIYMPKWTLL 979
DB 721 WWRNEQFVIGISAHLPFAVQGLKVLAGIDTNTFTVTSKANDERGDFAEIYMPKWTLL 780
QY 980 IPPTTILLINNVGVVAGTSYAINSNGSWGLFGKLPFAFWVIVHLYPFLKGLMGRQNT 1039
DB 980 IPPTTILLINNVGVVAGTSYAINSNGSWGLFGKLPFAFWVIVHLYPFLKGLMGRQNT 840

Db 781 IPPTTILLINNVGVVAGTSYAINSNGSWGLFGKLPFAFWVIVHLYPFLKGLMGRQNT 840
QY 1040 PTIVIVWAVLLASIFSLWVRDPFTRLAGPNQTCGNC 1080
DB 841 PTIVIVWAVLLASIFSLWVRDPFTRTRVGTQTCGNC 881

RESULT 4
US-09-221-013A-6
; Sequence 6, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Feng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-6

Query Match      68.6%; Score 3963.5; DB 4; Length 1081;
Best Local Similarity 68.0%; Pred. No. 0;
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;

QY 2 DGDALAKSGRHGAGDVCGIACDGLGTTLDGVFTACDVCRPCPCYEHREKGTQAC 61
DB 25 DGTGTPKLNMMN---CGICGIDGDDVGLAETGDFVACNECAPPCPCYEHREKGTQCC 81
QY 62 LQCKTKYKRRGSPAIRGEGDDTDADDGSDPNYPASGTEDQOKIADRMRSWRMTGGS 121
DB 82 PQCKTRFRHRGSPVEGEDDEDVDDIENEFYAGKANKARHQHGEFSS----- 133
QY 122 GNVHPKYSSEIGLSKYD---SGEIPRGYVPSVTNSQMSGEIPGASPDHMMSPGTNIS 178
DB 134 ----SSRHSQPIPLLTHTGHTVSGEIRTPDQSVRTT--SGPL-GPSDRNATSSP--YID 184
QY 179 PRAPPYVNHSPNPSREF-SGSIGNVAMKERVDMKMDKQKGAIPMTNGTSTAPSEGRA 237
DB 185 PRQPVV--RIVDFSKOLNSYGLGNVDWKERVGKWKQKQKNMLQMTG----KYHEGRGG 238
QY 238 TDIDASTEYNMEDALLNDETROPISRKVPPIASSKINPYRMVILRLVLSIFLHYRLTNP 297
DB 239 -EIE-GTSGNGEELQMADDTRLPMRSRVVPIFSSRLTPYRVVILLRLILCFLLQVTRTHP 296
QY 298 VRNAYPLMLLSVICEIFALSWILDQPPKWPPIRETLYDLRALAYDREGPSQLAAVDI 357
DB 297 VKNAYPLMLTSVICEIFALSWILDQPPKWPPIRETLYDLRALAYDREGPSQLVPVDV 356
QY 358 FVSTVDPLKEPPIVANTVLSILAVDPVKVSCVSDGASMLTTFDALAETSEFARKW 417
DB 357 FVSTVDPLKEPPIVANTVLSILAVDPVKVSCVSDGASMLTTFDALAETSEFARKW 416
QY 418 PFVKYDIEPRAPEYFQCKIDYLDKQVPSFVKDRRAMKREYEFKIRINALVSKALKV 477
DB 417 PFCKENLEPRAPEYFQCKIDYLDKQVPSFVKDRRAMKREYEFKIRINALVSKALKV 476
QY 478 PEEGWIMQDGTWPWGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVYVSREKRFQGHK 537
DB 477 PEEGWIMQDGTWPWGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVYVSREKRFQGHK 536
QY 538 KAGAMNALVRVSAVLNQGMYLNLDCHYINNSKAVREAMCFMDPNLGPVYQFQRFQ 597
```

Db 537 KAGAMNLRVSAVLNGLNGAYLLNVDCDHFNNSKAIKAMCFMMDPAIGKKCYVQFPOR 596
QY 598 FGDIDRNDYARNRTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 655
Db 597 FGDIDLHDYARNENIVFFDINMKGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 656
QY 656 FLA-SLCGGKKKASKKSSDKKKSNKHVDSSVPVFNLEDIBEGVEGAGFDDEKSVLMS 714
Db 657 IIVKSCGSRKKGSKKYNKRRGINRSDSNAPLFNMEDIDEGFE--GYDDERSILMS 714
QY 715 QMSLEKRFQGSAAFAVASTLMYEGVQSPSTPEILLKEAIIHVISCGYEDKSEWGTGIGWIY 774
Db 715 QRSVEKRFQGSAAFAVASTLMYEGVQSPSTPEILLKEAIIHVISCGYEDKSEWGTGIGWIY 774
QY 775 GSVTEDLITGFKMHARGWRSVYCMKRPAPFKGSAPIINLSRDLNQLRWALGSEVILFSRH 834
Db 775 GSVTEDLITGFKMHARGWRSVYCMKRPAPFKGSAPIINLSRDLNQLRWALGSEVILFSRH 834
QY 835 CPLWYGYGRLKFLERFAYINTTIVPITSPLIAYCILPAFCUITHRFIIPELSNYASIW 894
Db 835 CPLWYGYGRLKFLERFAYINTTIVPITSPLIAYCILPAFCUITHRFIIPELSNYASIW 894
QY 895 FIALFISIFATGILEMRWSGVIEDWNRNQFVWIGGISAHLFVAFQGLLKVLGIDTNF 954
Db 895 FIALFISIFATGILEMRWSGVIEDWNRNQFVWIGGISAHLFVAFQGLLKVLGIDTNF 954
QY 955 TVTISKANDEGDFAEILYFKNWKTLLIPPTTILINMVGVVAGTSYAINSGYQSWGFLPGK 1014
Db 955 TVTISKANDEGDFAEILYFKNWKTLLIPPTTILINMVGVVAGTSYAINSGYQSWGFLPGK 1014
QY 1015 LFFAFWVIAHLVPLKGLGRQNTPTTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1072
Db 1015 LFFAFWVIAHLVPLKGLGRQNTPTTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1072

RESULT 5

US-09-221-013A-12
; Sequence 12, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221.013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-12

Query Match 68.5%; Score 3959.5; DB 4; Length 1081;
Best Local Similarity 67.9%; Pred. No. 0;
Matches 732; Conservative 138; Mismatches 169; Indels 39; Gaps 15;

QY 2 DGDADALSGRHGADVCQICADGLGTLTLDGVFTACDVCRPPVCRPCYEHKEGTQAC 61
Db 25 DGGTKPLKNWN--GQICQICGDDVGLAETGDFVACNECAFPPVCRPCYEHKEGTQAC 81
QY 62 LQCKTKYKRRGSAIRGEEDDDTADGDSFNYPASTEDQOKIADRMSWRNMTGS 121
Db 82 PQCKTRFRHRGRSPVEGDEDDVDVDDIENEFNVAQGANRKARHQRHGEFFSS----- 133

QY 122 GNVGHPKYDSGEIGLSKYD---SGEIPRGYVSVTNSQMSGEIPGASPDHMMSPGTCNIS 178
Db 134 ---SRHESQPIPLLTHGTVSGEIRTPDTQSVRTT--SGPL-GPSDRNAISSP--YID 184
QY 179 RRAPPFYVNHSPNSREP--SGSIGNVAVKERVDCWQKODKGAIPMTNGTSTIAPSGRAA 237
Db 185 PRQFPVPV--RIVDPKSLNSYGLGNVDKERVGVKQKQKMLQWTG---KYHGGKGG 238
QY 238 TUIDASTEYNMEDALLNDETROPLSRKVPFIASSKINPVMVILRLVWLSIFLHYRLTNP 297
Db 239 -EIE-GTSGNGBEQWADDTRLPMSRVPIPSRLTPYRVVILRLILLCFFLOYRTHP 296
QY 298 VRNAYPLWLLSVCEIWFALSWILDQFPKWPINRETYLDRALRYDREGEPSQLAAVDI 357
Db 297 VKNAYPLWLLSVCEIWFALSWILDQFPKWPINRETYLDRALRYDREGEPSQLAAVDI 356
QY 358 FVSTVDPLEKPPVITANTVLSILAVDYPDKVSCYVDDGASMLTDAEAETSEFARKWV 417
Db 357 FVSTVDPLEKPPVITANTVLSILAVDYPDKVSCYVDDGASMLTDAEAETSEFARKWV 416
QY 418 PFVKYDIDIPRAPEFYFCQKIDYLDKVPSPFVKDRAMKREYEEFKIRINALVSKALKV 477
Db 417 PFCKKNIERAPREFYFAQKIDYLDKVPSPFVKDRAMKREYEEFKIRINALVSKALKV 476
QY 478 PEEGWIMQDGTWPNGNTRDHPQMIQVFLHSGGLDTEGNEPLRLVTVSREKPGFQHHK 537
Db 477 PEEGWIMQDGTWPNGNTRDHPQMIQVFLHSGGLDTEGNEPLRLVTVSREKPGFQHHK 536
QY 538 KAGAMNLRVSAVLNGLNGAYLLNVDCDHFNNSKAIKAMCFMMDPAIGKKCYVQFPOR 597
Db 537 KAGAMNLRVSAVLNGLNGAYLLNVDCDHFNNSKAIKAMCFMMDPAIGKKCYVQFPOR 596
QY 598 FGDIDRNDYARNRTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 655
Db 597 FGDIDLHDYARNENIVFFDINMKGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 656
QY 656 FLA-SLCGGKKKASKKSSDKKKSNKHVDSSVPVFNLEDIBEGVEGAGFDDEKSVLMS 714
Db 657 IIVKSCGSRKKGSKKYNKRRGINRSDSNAPLFNMEDIDEGFE--GYDDERSILMS 714
QY 715 QMSLEKRFQGSAAFAVASTLMYEGVQSPSTPEILLKEAIIHVISCGYEDKSEWGTGIGWIY 774
Db 715 QRSVEKRFQGSAAFAVASTLMYEGVQSPSTPEILLKEAIIHVISCGYEDKSEWGTGIGWIY 774
QY 775 GSVTEDLITGFKMHARGWRSVYCMKRPAPFKGSAPIINLSRDLNQLRWALGSEVILFSRH 834
Db 775 GSVTEDLITGFKMHARGWRSVYCMKRPAPFKGSAPIINLSRDLNQLRWALGSEVILFSRH 834
QY 835 CPLWYGYGRLKFLERFAYINTTIVPITSPLIAYCILPAFCUITHRFIIPELSNYASIW 894
Db 835 CPLWYGYGRLKFLERFAYINTTIVPITSPLIAYCILPAFCUITHRFIIPELSNYASIW 894
QY 895 FIALFISIFATGILEMRWSGVIEDWNRNQFVWIGGISAHLFVAFQGLLKVLGIDTNF 954
Db 895 FIALFISIFATGILEMRWSGVIEDWNRNQFVWIGGISAHLFVAFQGLLKVLGIDTNF 954
QY 955 TVTISKANDEGDFAEILYFKNWKTLLIPPTTILINMVGVVAGTSYAINSGYQSWGFLPGK 1014
Db 955 TVTISKANDEGDFAEILYFKNWKTLLIPPTTILINMVGVVAGTSYAINSGYQSWGFLPGK 1014
QY 1015 LFFAFWVIAHLVPLKGLGRQNTPTTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1072
Db 1015 LFFAFWVIAHLVPLKGLGRQNTPTTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1070

RESULT 6

US-09-221-013A-8
; Sequence 8, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.

QY	760	YEDXSEWGTEIGWYGSVTE	ILTFGKMHARGWSVYCM	PKRPAFKGSAPINLS	DRLNQV	811
Db						
QY	764	YODKTEWGKEIGWYGSVTE	ILTFGKMHCHGWSVYCM	PKRAAFKGSAPINLS	DRLHQV	823
Db						
QY	820	LRWALGSVEILFSRHCP	LWYGGGRLLKFLRPA	YNTTIYPLTSLPLLVY	CILPAICLLT	879
Db						
QY	824	LRWALGSVEIFLSRHCP	IWIYGGGLKWLRFVS	INSVYVPTWSLPTVY	CSLPAVCILLT	883
Db						
QY	880	GKFIPEISNLASITWFI	ALFLSIFATGILEMR	WSGVGIDENWRNE	OFWVIGGISAHLPAV	939
Db						
QY	884	GKFIPEISNVAGILF	MDMFISIAVTGILEM	QGVGIDDDWRNEQ	FWVIGGASSHLPA	943
Db						
QY	940	FQGLLKVLAGIDTNT	FTVTSKANDEEGDFA	EYLMFKWITLLIP	PTTLLIINNMGVVAGTSY	999
Db						
QY	944	FQGLLKVLAGVNTNT	FTSKAAD-DGAFSE	LYIFKWITLLIP	PTTLLIINIIGVVGSD	1002
Db						
QY	1000	AINSYGWSGBLFK	GLFPAFWVHLVYPL	KGLMGRQNRPT	TVIVWAVLLASIFSLWV	1059
Db						
QY	1003	AISNGYDSWGLF	GRGLFFALWVHLV	YPLFKMLGKODK	MPITIIIVWSILASITLWV	1062
Db						
QY	1060	RVPDPFTRLAGE	NIQTGCINC	1080		
Db						
QY	1063	RINFVAK-GGPV	LEICLNC	1082		
Db						
RESULT 7						
US-08-960-048-6						
; Sequence 6, Application US/08960048C						
; Patent No. 6271443						
; GENERAL INFORMATION:						
; APPLICANT: Stalker, D. et al.						
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter						
; TITLE OF INVENTION: Sequences						
; FILE REFERENCE: 15621/01/US						
; CURRENT APPLICATION NUMBER: US/08/960,048C						
; CURRENT FILING DATE: 1997-10-29						
; PRIOR APPLICATION NUMBER: 60/029,987						
; PRIOR FILING DATE: 1996-10-29						
; NUMBER OF SEQ ID NOS: 12						
; SOFTWARE: FastSeq for Windows Version 4.0						
; SEQ ID NO 6						
; LENGTH: 974						
; TYPE: PRT						
; ORGANISM: Gossypim hirsutum						
US-08-960-048-6						
Query Match 60.2%; Score 3477; DB 3; Length 974;						
Best Local Similarity 61.5%; Pred.No.1.5e-318;						
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;						
QY	18	VCQICADGLGTLLDGDV	TACDVCRRPVCPCYEH	EKEGCTOACLOCKTKYK	KHRGSPAI	77
Db						
QY	8	VCHTCGHEHGLNVNG	BFFVACHECNFPICKS	CFEYDLKEGRKACLR-		58
Db						
QY	78	RGBEGDDTDADGSD	ENFPYASGTEDQKQI	ADRMRSWRMNTGSGN	GV-HPKYDSDGEIGL	136
Db						
QY	59	-----DENLLDD-	-----VEKATGDOST	MAA-----HLNKSQ	DVGIIHARHIS--SV	96
Db						
QY	137	SKYDSGEIPRGY	VPVSTVNSQMSGEI	PCASPDHMMGPTGN	ISRRAPFPVYVNHSP	NSREF 196
Db						
QY	97	STLDS-----				EM 103
Db						
QY	197	SGSIGNVANKERV	DGWMKODKGAIP	WTNGTSTIAPSE	GRAATDIDASTEY	NMEDALLNDE 256
Db						
QY	104	AENGNSIWKNR	VESWEKKNKKKK	PATT-----KVERAE	EIPPEQOMDEK	PAPDA 154
Db						
QY	257	TROPLSRKVP	PIASSKINPYRMV	ITLVRLVLSIFL	HYRLTNPNVRYN	AYPLMLLSVICIWEFA 316
Db						
QY	155	S-QPLSTI	PIPKSLAPRYTV	IIMELIILGLF	FHYRVINPVD	SAGFLWLTSVICIWEFA 213
Db						
QY	317	LSWILDOFP	KWFPINRETYL	DRALRYDREG	EPSLAAVDI	FVSTVDPLKEPPIVANTV 376
Db						
QY	214	FSWVLQFP	KWYVNPRTYI	DRLSARYEREG	PDALAAVDF	FVSTVDPLKEPPIVANTV 273
Db						

```

QY 377 LSTILAVDPVDKVCVSDGASMLTFDALAETSEFARKWVPVKKYDIEPRAPFVFCQ 436
Db 274 LSTILADVPVDKVCVSDGASMLTFESLVEFADFAKWKVPCKFESIEPRAPFVFSQ 333
QY 437 KIDYLDKQVPSFVKORRAMKREYEFKIRINALVSKALKVPBEGMTDGTWPNGNTR 496
Db 334 KIDYLDKQVPSFVKERRAMKRDYEVKIRINALVAKAQTPDEGTMQDGTWPNGNPR 393
QY 497 DHPGMIQVFLHSGGLDTEGNELPRLVYVSREKPGFOHKKAGAMNALVRVSAVLNQG 556
Db 394 DHPGMIQVFLGSGARDIEGNELPRLVYVSREKPGYQHHKAGAMNALVRVSAVLNAP 453
QY 557 YMLNLCDDHYNNKSKAVREAMCFMDPNLGPQVCYVQFQORFDGIDRNDRYANRNTVFFD 616
Db 454 FIUNLDCDDHYNNKSKAVREAMCFMDPQVGRDVCYVQFQORFDGIDRNDRYANRNTVFFD 513
QY 617 INLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPITAKKPGFLASLGGKKKASKSKRSSD 676
Db 514 VNMKGLDGIQGPVYVGTGCVFNRAQALYGYGPPSNPSPKSSSSSSSCSCCCPGKPKDPSE 573
QY 677 KKSNNKHVDSSVPVFNLEDEEGVEGAGFDD-EKSVLMSQMSLEKRFQSOAAFAVASTLME 735
Db 574 LYRDAKREELDAAIFNLRREID-----NYDEYERSMLISQTSFEKTFGLSSVFIESTLME 627
QY 736 YGVVQSSPTPESLLKEAIIHVISCGYEDKSEWGTGTEIGWYGSVTEIDILTGFKMHARGWRSV 795
Db 628 NGGVAESANPSTLIIKEAIIHVISCGYEBKTANGKEIGWYGSVTEIDILTGFKMHARGWRSI 687
QY 796 YCMKPRPAFGKSAPINLSRDLNOVLRWALGSVEILFSRHCPPLWYGY-GGRILKFLERPAYI 854
Db 688 YCMPLRPAFGKSAPINLSRDLHQLVLRWALGSVEILFSRHCPPLWYGYGGGRLKWLQRLAYI 747
QY 855 NTIYPLTSLPLVYCYLTPALCILLTGKFIEMPEISNLASIWFIALFLSIFATGILEMWSG 914
Db 748 NTIYPTSLPLAYCSLPAICLLTGKFIPTLSNLASVILFLGLFSLIIVTAVLELRWSG 807
QY 915 VGDEWWRNEQFWVIGGISAHFLFAVFOGLLKLVLAGIDTNTFTVTSKANDEEGDFAELMYFK 974
Db 808 VSIEDLWRNEQFWVIGGVSAILFAVFOGLKMLAGIDTNTFTVTAKAAD-DADFGELYIVK 866
QY 975 WTTLLIPPTTILIIINMVGVVAGTSYAINSGYQSGWPLFGKLPFAFWIVHLYPFLKGLMG 1034
Db 867 WTTLLIPPTTLLIINMVGVVAGFSDALNKGYEAWGPLFGKVFVFSFWILHLYPFLKGLMG 926
QY 1035 RQNRTPITIVWAVILLASIFSLWVRVDPPTTRLAGPNI-QTC-GINC 1080
Db 927 RQNRTPITIVWSVLLASVFSVLWVRINPFPVSTADSTTVSQSCLSIDC 974

```

RESULT 8

```

US-09-838-586-6
; Sequence 6, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-838-586-6

```

```

Query Match 60.2%; Score 3477; DB 4; Length 974;
Best Local Similarity 61.5%; Pred. No. 1.5e-318;
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;

QY 18 VCOICADGLGTLTLDGVPFTACDVCPPVCRPCVEHERKEGTQACLOCKTKYKRRGSPAI 77
Db 8 VCHTCGEHVLNVNGBFFVACHCECNFPIKSCPEYDLKEGRKACLR-----GSPY- 58
QY 78 RGEGBGDDTADGSDPNYPASGTEDQKQIADRMRSWRMNTGGSGNVG-HPKYDSIGL 136
Db 59 -----DENLDD-----VEKATGDOSTMAA-----HLNKSQDVIHARHIS--SV 96
QY 137 SKYDSGEIPRGYVPSVTNSQMSGEIFGASPDHMMSPGTGNI SRRAPFPYVNHSPNREF 196
Db 97 STLDS-----EM 103
QY 197 SGSIGNVAMKERVDMKMKODKGAIPTWNTGTSIAPSEGRAATDIDASTEYNNMEDALINDE 256
Db 104 AEONGNSIMKNRVSVESWEKKNKKKPAATT-----KVERAEIIPPEQOMEDKAPDA 154
QY 257 TROPLSRKVPPIASSKINPMVIVLRLVLSIFELHYRLTNPNVBNAYPLMLLSVICIWFPA 316
Db 155 S-QPLSTIIPKSRAPYRTVIIMRLIILGLFFHYRTNPNVDSAFGLMLTSLVICIWFPA 213
QY 317 LSWILDQFPKWFPIINRETYLDRALRYDREGEPSQLAANDIFVSTVDPLKEPPIVTANTV 376
Db 214 FSWVLQFPKWPVNVNRETVIDLSARYEREGETDELAANDFFVSTVDPLKEPPIVTANTV 273
QY 377 LSTILAVDPVDKVCVSDGASMLTFDALAETSEFARKWVPVKKYDIEPRAPFVFCQ 436
Db 274 LSTILADVPVDKVCVSDGASMLTFESLVEFADFAKWKVPCKFESIEPRAPFVFSQ 333
QY 437 KIDYLDKQVPSFVKORRAMKREYEFKIRINALVSKALKVPBEGMTDGTWPNGNTR 496
Db 334 KIDYLDKQVPSFVKERRAMKRDYEVKIRINALVAKAQTPDEGTMQDGTWPNGNPR 393
QY 497 DHPGMIQVFLHSGGLDTEGNELPRLVYVSREKPGFOHKKAGAMNALVRVSAVLNQG 556
Db 394 DHPGMIQVFLGSGARDIEGNELPRLVYVSREKPGYQHHKAGAMNALVRVSAVLNAP 453
QY 557 YMLNLCDDHYNNKSKAVREAMCFMDPNLGPQVCYVQFQORFDGIDRNDRYANRNTVFFD 616
Db 454 FIUNLDCDDHYNNKSKAVREAMCFMDPQVGRDVCYVQFQORFDGIDRNDRYANRNTVFFD 513
QY 617 INLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPITAKKPGFLASLGGKKKASKSKRSSD 676
Db 514 VNMKGLDGIQGPVYVGTGCVFNRAQALYGYGPPSNPSPKSSSSSSSCSCCCPGKPKDPSE 573
QY 677 KKSNNKHVDSSVPVFNLEDEEGVEGAGFDD-EKSVLMSQMSLEKRFQSOAAFAVASTLME 735
Db 574 LYRDAKREELDAAIFNLRREID-----NYDEYERSMLISQTSFEKTFGLSSVFIESTLME 627
QY 736 YGVVQSSPTPESLLKEAIIHVISCGYEDKSEWGTGTEIGWYGSVTEIDILTGFKMHARGWRSV 795
Db 628 NGGVAESANPSTLIIKEAIIHVISCGYEBKTANGKEIGWYGSVTEIDILTGFKMHARGWRSI 687
QY 796 YCMKPRPAFGKSAPINLSRDLNOVLRWALGSVEILFSRHCPPLWYGY-GGRILKFLERPAYI 854
Db 688 YCMPLRPAFGKSAPINLSRDLHQLVLRWALGSVEILFSRHCPPLWYGYGGGRLKWLQRLAYI 747
QY 855 NTIYPLTSLPLVYCYLTPALCILLTGKFIEMPEISNLASIWFIALFLSIFATGILEMWSG 914
Db 748 NTIYPTSLPLAYCSLPAICLLTGKFIPTLSNLASVILFLGLFSLIIVTAVLELRWSG 807
QY 915 VGDEWWRNEQFWVIGGISAHFLFAVFOGLLKLVLAGIDTNTFTVTSKANDEEGDFAELMYFK 974
Db 808 VSIEDLWRNEQFWVIGGVSAILFAVFOGLKMLAGIDTNTFTVTAKAAD-DADFGELYIVK 866
QY 975 WTTLLIPPTTILIIINMVGVVAGTSYAINSGYQSGWPLFGKLPFAFWIVHLYPFLKGLMG 1034
Db 867 WTTLLIPPTTLLIINMVGVVAGFSDALNKGYEAWGPLFGKVFVFSFWILHLYPFLKGLMG 926
QY 1035 RQNRTPITIVWAVILLASIFSLWVRVDPPTTRLAGPNI-QTC-GINC 1080

```


Db 598 GVVAGVSDAINNGYSGWGLFGLKLFAPFWIHLHLPFLKGLMGRQRTTIVVLSILLA 657
QY 1052 SIFSLLWVRVDPPTTRLAGENIQTGCI 1078
Db 658 SIFSLLWVRVDPPTTRLAGENIQTGCI 684

RESULT 11
US-09-221-013A-2
; Sequence 2, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-2

Query Match 45.5%; Score 2629.5; DB 4; Length 629;
Best Local Similarity 75.2%; Pred. No. 4.6e-239;
Matches 468; Conservative 76; Mismatches 71; Indels 7; Gaps 4;

QY 454 RAMKREYEEFKIRINALVSKALVPERGWINODGTPMGNTRDHPGMIQVFLHSGGLD 513
Db 1 RAMKREYEEFKIRINALVSKALVPERGWINODGTPMGNTRDHPGMIQVFLHSGGLD 60
QY 514 TEGNELPRLVVSREKRGPFQHHKKAGAMNALVRVAVLTNGQYMLNLCDDHYINNSKAV 573
Db 61 TDGNEPRLIVVSREKRGPFQHHKKAGAMNARSIRVAVLTNGAYLLNVDCDHYENNSKAI 120
QY 574 REAMCFMDPRLGQVQVQPPQFPGIDRNDVANNRVFPDINRLGLDGIQGVVYGT 633
Db 121 KEAMCFMMDPAIGKKCCVQVQPPQFPGIDRNDVANNRVFPDINRLGLDGIHGPVYGT 180
QY 634 GCVENRTAIYGEPIKAK--KPGFLA--SLCGGKKKASKSKRSSDKKSNKHVDSSVPV 690
Db 181 GCCFNQALYGYDPLVTEEDLEPNLIVKSCGSRKKGSKSKYVEKRGINRSDSNAPL 240
QY 691 FNLEIDIEGVGAGFDDEKSVLMSQMSLEKRFQSGAFAVSTLMEYGGVQSSPTESLLK 750
Db 241 FNMEDIDGFE--GYDDERSILMSQSVKEKFGQSPVFIATFMEQGGIPPTNPATLLK 298
QY 751 EAIHVISCGYEDKEMGWTEIGWYGSVTEIDILTGFMHARGWRSVYCMKPRPAFGSAPI 810
Db 299 EAIHVISCGYEDKEMGWTEIGWYGSVTEIDILTGFMHARGWRSVYCMKPRPAFGSAPI 358
QY 811 NLSRLNQLRWALGSGVILSRHCPVWYGGRLKFLERFAYINTTIYPLTSLPLLYVC 870
Db 359 NLSRLNQLRWALGSGVILSRHCPVWYGGRLKFLERFAYINTTIYPLTSLPLLYVC 418
QY 871 ILPAICLITGTFIMPEISINLASIFALFISFATGILEMRWSGVGIDENWRNEQFWVIG 930
Db 419 ILPAICLITGTFIMPEISINLASIFALFISFATGILEMRWSGVGIDENWRNEQFWVIG 478
QY 931 GISAHFLAVFQGLLKVLAGIDNFTVTSKANDDEGDAELVYFKWTTLLIPTTILLINM 990
Db 479 GTSTHFLAVFQGLLKVLAGIDNFTVTSKANDDEGDAELVYFKWTTLLIPTTILLINM 538

QY 991 VGVAGTGYAINNGYSGWGLFGLKLFAPFWIHLHLPFLKGLMGRQRTTIVVWVLL 1050
Db 539 IGIVAGSVANNNGYSGWGLFGLKLFAPFWIHLHLPFLKGLMGRQRTTIVVWVLL 598
QY 1051 ASIFSLLWVRVDPPTTRLAGPN 1072
Db 599 ASIFSLLWVRVDPPTTRLAGPN 618

RESULT 12
US-09-221-013A-14
; Sequence 14, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-221-013A-14

Query Match 30.4%; Score 1756.5; DB 4; Length 547;
Best Local Similarity 61.5%; Pred. No. 1e-156;
Matches 346; Conservative 66; Mismatches 100; Indels 51; Gaps 13;

QY 2 DGDA-DALKSGRHGAGDVQCICADGLTTLGDVFTACDVCRFPVPCPYEHEKGTQA 60
Db 23 DGDAPPKAPGKSVNGVQCICGDTGVGSATGDVFAVACNECAFVCPVCPYERKEGNC 82
QY 61 CLQCKTKYKHRGSPAIRGEGDDTDADD--GSDPNP---ASGTEDQKQIADMRMRWM 116
Db 83 CPQCKTKYKHRGSPAIRGEGDDTDADD--GSDPNP---ASGTEDQKQIADMRMRWM 129
QY 117 NTGSGNVGHPKYDSGEIGLSKYDSGEIPRGPVPSVTNSQMSGEIPGASPDHMMSPGN 176
Db 130 QRQGE-----DVLSSSRHEQHRIPR---LTSQOQISGEIPDASPDHHSI--- 172
QY 177 ISRRAPPFYVNHSS-----PNPSREF--SGSIGNVAMKRVGDKMKQDKGAIPTWNGTS 228
Db 173 --RSGTSYVDPSVPVFPVRIVDPSKDLNSYSGINSVDMQERVASWRNKQDKNMQVAN--- 227
QY 229 IAPSEGAATDIDASTYNNMEDALLNDETROPKSRKVPKPIASSKINPRMVIVLRLVLSI 288
Db 228 -KYPEARGG-DME-GTSGNGEDIQWDDARLPLSRIVIPSNQNLNRYIVILILILIMF 284
QY 289 FLHRLTNPRVNAVPLMLLSVICIWFALSMLDQFQKWFPPINRETYLDRLALRYDREG 348
Db 285 FQFRVTHPVADATGLMLVSVICELWPLSLWLLQFQKWFPPINRETYLDRLALRYDREG 344
QY 349 PSQAALVDIFVSTVDPLKEPPIVNTVLSILAVDVPDVKVSCYVSDDGASMLTFDALAE 408
Db 345 PSQAALVDIFVSTVDPLKEPPIVNTVLSILAVDVPDVKVSCYVSDDGASMLTFDALAE 404
QY 409 TSEPAKRVVFKKVDIEPRAPFEVFCOKIDYLDKQVPSFVKDRAMKREYEEFKIRIN 468
Db 405 TAEFAKRVVFKKVDIEPRAPFEVFCOKIDYLDKQVPSFVKDRAMKREYEEFKIRIN 464
QY 469 ALVSKALVPBEGTMDGTTPWGNTRDHPGMIQVFLHSGGLDTEGNEPRLVVSRE 528
Db 465 ALVSKALVPBEGTMDGTTPWGNTRDHPGMIQVFLHSGGLDTEGNEPRLVVSRE 524

QY 529 KRPGFOHHKAGAMNALVRVSAV 551
Db 525 KRPGFOHHKAGAMNALIRVSAV 547

RESULT 13

US-08-812-008-48
; Sequence 48, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-812-008-48

Query Match 4.9%; Score 282; DB 4; Length 57;
Best Local Similarity 92.9%; Pred. No. 3.8e-19;
Matches 52; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 774 YGSVTEDLTGKMHARGWRSVYCMKRPAPKGSAPINLSDRLNQVLRWALGSGVEI 829
Db 1 YGSVTEDLTGKMHARGWRSVYCMKRPAPKGSAPINLSDRLNQVLRWALGSGVEI 56

RESULT 14

US-08-960-048-11
; Sequence 11, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-960-048-11

Query Match 4.6%; Score 268; DB 3; Length 693;

Best Local Similarity 19.5%; Pred. No. 5.7e-16;
Matches 156; Conservative 83; Mismatches 201; Indels 360; Gaps 29;

QY 263 RKVPIASSKINPYRMVILRLVLSIFLHYRUT-----NPVRNAYPLWLLSVICIEWFA 316
Db 11 RRMGRFSAL-----MLIVLSLTVSCRYIWMRYTSTLNWDDPVSIVGLLILFAITYAVIV 66
QY 317 LSWILDQPKWPPINRETYLDRLALRYDEGSPQLAADVIFVSTVDPKPEPIVYANTV 376
Db 67 L--VLGYFQVWPLNRQP-----VPLPKDMSLWPS---VDIFVPTYN---EDLVVVKNTI 113
QY 377 LSILADVDPVKVSCVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFPYFCQ 436
Db 114 YASLGIDWPCKLNIWLDGG----- 135
QY 437 KIDYLDKQVQPSFVKDRAMKREYEEFKIRINALVSKALKVPEEGWIMQDGTWPGNNTR 496
Db 136 -----RREFRQFAQNVG----- 147
QY 497 DHPGMIQVFLGHSGLDTGNEPLPLVYVSRKRGFQHHKAGAMNALVRYSAVLTNQ 556
Db 148 -----VKYIARTT-----HEHAKAGINNALKVA---KGE 174
QY 557 YMLNLDCDHYINNSKAVREAMCFMLDPNLGPOVCYVQPPQRE---DGIDRN---DRYAN 609
Db 175 FVSIFDCDHVPTSRFLQMTMGWFLKE---KQLAMQTHHFFSDPFPERNLGRFRKTEN 230
QY 610 RNTVFEDINLRGLDIGFVYVGTGCVFNRTAIYGYEPPIKAKKFGFLASLGGKKKASK 669
Db 231 EGTLYFGLVQDGNMDWDATFFCGSCAVIRR-----KE----- 262
QY 670 SKKRSSDKKSNKHVDVSVVPFNLEDIEGVEGAGFDDKSVLMSQMSLEKRFQGSAAV 729
Db 263 ----- 262
QY 730 ASTLMEYGGVPQSSTPESLLKEAIIHVISGVEDKSEWGTGTEIGWYGSVTEDLTGKMH 789
Db 263 ---LDEIGGI-----AVE-----TVTEDAHTSLRHR 286
QY 790 RGWRSVYCMKRPAPKGSAPINLSDRLNQVLRWALGSGVEILFSRHCPHLYWYGGSLKFL 849
Db 287 RGYTSAYM--RIPQAAGLATESLSAHIQIRIRWARGMVQI--FRLDNPL---TGKGLKFAQ 340
QY 850 RFAYINTIYPLTSLPLVAVCILP-----AICLTGKFIIMP----- 885
Db 341 RLCVYNAHFELSGIPRLIFLTAPLAFLLHAYIIYAPALMIALFVLPFMIHSLTNSKI 400
QY 886 -----EISNLASIWFTIA---LFLSIFATGILEMRWSGVGIDSWRNEOF--WVI 929
Db 401 QGKVRHSFWSSEIYETVLAWYIAPPTLVALINPHKGFKNVTAKGGGLVE---EEVDWVI 456
QY 930 GGISAHILFAVFQGLLKVLAGIDTNTVTSKANDEBGFPAELMFKWTTLIPPTTIL--I 987
Db 457 S--RPYIFPLVLLNLVGVAVGI-----WRYFYGPTEMLTVV 490
QY 988 INWGV-----VAGTSYAIN 1002
Db 491 VSMVWVFNVLIVLGGAVAVS 510

RESULT 15
US-09-838-586-11

```
; Sequence 11, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-838-586-11
```

```
Query Match          4.6%; Score 268; DB 4; Length 693;
Best Local Similarity 19.5%; Pred. No. 5.7e-16;
Matches 156; Conservative 83; Mismatches 201; Indels 360; Gaps 29;

QY 263 RKVPIASSKINPYRMVILRLVLSIFLHVRLT-----NPVRNAVPLMLLSVCEIWFA 316
Db 11 RRMPEGRFSA-----MLIVLSLTSCRYIWKRYTSTLWDDPVSIVCGLLILFALTAYMIV 66

QY 317 LSWILQDFPKWFINRETYLDRLALRYDREGPEPSQLAAVDIFVSTVDPLKEPPIVTANTV 376
Db 67 L--VLGYFQVWELNRQP-----VPLPKDMSLWFS---VDIFVPTYN---EDLNVVKNTI 113

QY 377 LSLIANDYPVDKSVSCYVDDGASMLTFDALAESEFARKWVPFVKKYDIEPRAPEFYFCQ 436
Db 114 YASLIGIDWPDKLNIWILDGG-----135

QY 437 KIDYLDKQVQSFVKDRAMKREVEEFKIRINALVSKALVPBEGIMQDGTWPWGNTR 496
Db 136 -----REFRQFAQNVG-----147

QY 497 DHPGMIQVFLGHSGGLDTEGNELPRLVYVSRKRPQGHKKAGAMNALVRVSAVLNQG 556
Db 148 -----VKYIARTT-----HEHAKAGINNALKYA---KGE 174

QY 557 YMLNLCDDHYNNKSAVREANCFMLDPNLGPQVCYVQFPORF---DGIDEN---DRYAN 609
Db 175 FVSIFDCDHVPTRSFLQMTGWFLKE---KQLAMMOTPHHFFSPDPFERNLGRFRKTPN 230

QY 610 RNTVFFDINLRGLDGIQGPVVVGCVFENETAIYGEPPIKAKKPGFLASLCGGKKKASK 669
Db 231 EGTILFYGLVQDGNMWDATFCGSCAVIRR-----KP-----262

QY 670 SKRSDKKKSNKHVDSSVPVFNLEDIERGVEGAGFDDKSVLMSQMSLEKRFQGSAAFV 729
Db 263 -----262

QY 730 ASTIMEYGVYQOSTPSSLKEATHVISCYVEDKSEWGTETGIWYGSVTEDLITGFKMHA 789
Db 263 ---LDEITGGI-----AVE-----TVTEDAHTSLRLHR 286

QY 790 RGMRSVYCMKPKRPAPKGSAPINLSDRLNQVLRWALGSVEILFSRHCPLMWYGGRLKPLE 849
Db 287 RGYTSAYM--RIPQAAGLATESLSAHICQIRWARGWQI--FRLDNPL---TGKGLAFQAQ 340

QY 850 RFAYINTTYPLTSLPLVYCILP-----AICLLTGKFIIMP-----885
Db 341 RLCYVNAMFHELSGIPRLIFLTAFLAFLLHAYIIYAPALMIALFVLPHMIHASLITNSKI 400

QY 886 -----RISNLASTWFA---LFLSIFATGILEMRWSGVGIDBWNENQF--WVI 929
Db 401 QGKYRHSPWSBIYETVLAWIYAPPTLVALINPHKGFNVTAKGGLVE-----EYVDWVI 456
```

```
Search completed: August 23, 2004, 01:11:16
Job time : 29 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 01:09:20 ; Search time 93 Seconds
(without alignments)

3649.414 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALXSGRHGAGDVQC.....VDPFTTLRAGNIQTGGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1295152 segs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5778	100.0	1080	9	US-09-900-237-30
2	5452.5	94.4	1081	16	US-10-437-963-174885
3	5423.5	93.9	1079	12	US-10-627-132-22
4	5423.5	93.9	1079	14	US-10-267-459-6
5	5423.5	93.9	1079	14	US-10-209-059-22
6	5421.5	93.8	1077	12	US-10-627-132-10
7	5421.5	93.8	1077	14	US-10-209-059-10
8	5421.5	93.8	1077	14	US-10-160-719-6
9	5421.5	93.8	1077	14	US-10-160-719-30
10	5261.5	91.1	1043	14	US-10-160-719-50
11	5149	89.1	1076	12	US-10-627-132-14
12	5149	89.1	1076	14	US-10-209-059-14
13	5149	89.1	1076	14	US-10-160-719-58
14	5129	88.8	1119	12	US-10-425-114-58605
15	5122.5	88.7	1063	16	US-10-437-963-142245
					Sequence 30, Appl
					Sequence 174885,
					Sequence 22, Appl
					Sequence 6, Appli
					Sequence 22, Appl
					Sequence 10, Appl
					Sequence 6, Appli
					Sequence 30, Appl
					Sequence 50, Appl
					Sequence 14, Appl
					Sequence 58, Appl
					Sequence 58605, A
					Sequence 142245,

16	4507	78.0	1065	9	US-09-900-237-33	Sequence 33, Appl
17	4499	77.9	1065	14	US-10-229-193-10	Sequence 10, Appl
18	4455.5	77.1	881	9	US-09-838-539-8	Sequence 8, Appli
19	4031.5	69.8	1085	12	US-10-424-599-249738	Sequence 249738
20	4021.5	69.6	1091	9	US-09-900-237-26	Sequence 26, Appl
21	3993.5	69.1	1074	12	US-10-627-132-46	Sequence 46, Appl
22	3993.5	69.1	1074	14	US-10-209-059-46	Sequence 46, Appl
23	3993.5	69.1	1074	14	US-10-160-719-14	Sequence 14, Appl
24	3993.5	69.1	1074	14	US-10-160-719-22	Sequence 22, Appl
25	3993.5	69.1	1074	14	US-10-160-719-42	Sequence 42, Appl
26	3988	69.0	1075	12	US-10-627-132-2	Sequence 2, Appli
27	3988	69.0	1075	14	US-10-209-059-2	Sequence 2, Appli
28	3988	69.0	1075	14	US-10-160-719-10	Sequence 10, Appl
29	3988	69.0	1075	14	US-10-160-719-34	Sequence 34, Appl
30	3988	69.0	1075	14	US-10-160-719-54	Sequence 54, Appl
31	3963.5	68.6	1081	14	US-10-229-193-6	Sequence 6, Appli
32	3959.5	68.5	1081	14	US-10-229-193-12	Sequence 12, Appl
33	3890	67.3	1052	12	US-10-627-132-30	Sequence 30, Appl
34	3870.5	67.0	1055	16	US-10-437-963-117576	Sequence 117576,
35	3844.5	66.5	1094	12	US-10-627-132-18	Sequence 18, Appl
36	3844.5	66.5	1094	14	US-10-209-059-18	Sequence 18, Appl
37	3844.5	66.5	1094	14	US-10-160-719-26	Sequence 26, Appl
38	3844.5	66.5	1094	14	US-10-160-719-46	Sequence 46, Appl
39	3844.5	66.5	1165	9	US-09-900-237-8	Sequence 8, Appli
40	3830	66.3	847	12	US-10-424-599-238832	Sequence 238832,
41	3820.5	66.1	1039	9	US-09-900-237-14	Sequence 14, Appl
42	3820	66.1	1059	12	US-10-627-132-42	Sequence 42, Appl
43	3820	66.1	1059	14	US-10-209-059-42	Sequence 42, Appl
44	3820	66.1	1059	14	US-10-160-719-2	Sequence 2, Appli
45	3818	66.1	1086	9	US-09-900-237-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-900-237-30
; Sequence 30, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 1080
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-900-237-30

Query Match 100.0%; Score 5778; DB 9; Length 1080;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDGDADALXSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRCFPVCRPCYHERKEGQA	60
Db	1	MDGDADALXSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRCFPVCRPCYHERKEGQA	60
Qy	61	CLOCKTKYKRRHSGPAIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTGG	120
Db	61	CLOCKTKYKRRHSGPAIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTGG	120
Qy	121	SGNVGHPKYSGEIGLSKYSGEIPRGYVPSVTSQMSGEIPGASPDHMMSPGTNISR	180

Db 121 SGNVGHKPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMGFTGNISRR 180
QY 181 APPYVNHSPNPSRFSGSGIGNVANKERVGDGKWKQDKGAIPMTNGTSTIAPSEGRAATDI 240
Db 181 APPYVNHSPNPSRFSGSGIGNVANKERVGDGKWKQDKGAIPMTNGTSTIAPSEGRAATDI 240
QY 241 DASTEYNMEDALLNDETQPLSRKVPPIASSKNPYRMWIVLRVLVLSIFLHLYRLTNPNVN 300
Db 241 DASTEYNMEDALLNDETQPLSRKVPPIASSKNPYRMWIVLRVLVLSIFLHLYRLTNPNVN 300
QY 301 AYLWLLSVICETWFAWLSWILDQPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
Db 301 AYLWLLSVICETWFAWLSWILDQPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
QY 421 KKYDIEPRAPPEFYFCOKIDYLDKQVOPS FVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 421 KKYDIEPRAPPEFYFCOKIDYLDKQVOPS FVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
QY 481 GWTMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVREKRPFGQHHKKAG 540
Db 481 GWTMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVREKRPFGQHHKKAG 540
QY 541 ANNALVRVSAVLNTQYMLNLDCHYINNSKAVREACFLMDNLPQOVVCYVPPQPFDFG 600
Db 541 ANNALVRVSAVLNTQYMLNLDCHYINNSKAVREACFLMDNLPQOVVCYVPPQPFDFG 600
QY 601 IDRNDRVANRNTVFFDINLRGLDGIQGPVYVGTGCVENRTAIYGYEPPIKAKKPGFLASL 660
Db 601 IDRNDRVANRNTVFFDINLRGLDGIQGPVYVGTGCVENRTAIYGYEPPIKAKKPGFLASL 660
QY 661 CGGKKKASKSKRSDDKKSNKHVDSSVPVFNLEDEIEEGVEGAGDFDEKSVLMSQMSLEK 720
Db 661 CGGKKKASKSKRSDDKKSNKHVDSSVPVFNLEDEIEEGVEGAGDFDEKSVLMSQMSLEK 720
QY 721 RFGQSAAFVASTLMYEGVGPQSSPTPESSLKKEAIIHVISCGYEDKSEWGTGWIYGSVTE 780
Db 721 RFGQSAAFVASTLMYEGVGPQSSPTPESSLKKEAIIHVISCGYEDKSEWGTGWIYGSVTE 780
QY 781 ILTGFKMHARGWRVYCMKRPAPFKGAPINLSRLNQVLRWALGSVEILLFSRHCPLWYG 840
Db 781 ILTGFKMHARGWRVYCMKRPAPFKGAPINLSRLNQVLRWALGSVEILLFSRHCPLWYG 840
QY 841 YGGRKLFERPAYINTTIYPLTSLPLVYCILPAICLLTGKFIEMPEISNLASIWFIALFL 900
Db 841 YGGRKLFERPAYINTTIYPLTSLPLVYCILPAICLLTGKFIEMPEISNLASIWFIALFL 900
QY 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFVAVFQGLLKVLAGIDTNTFTVTSKA 960
Db 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFVAVFQGLLKVLAGIDTNTFTVTSKA 960
QY 961 NDEBGDPAELWFKWKTLLIPPTTILINMWVAVAGTSYAINSGYQSWGPLFGKLFFAFW 1020
Db 961 NDEBGDPAELWFKWKTLLIPPTTILINMWVAVAGTSYAINSGYQSWGPLFGKLFFAFW 1020
QY 1021 VIVHLYPFLKGLMGQRNTPPTIVIVWVLLASIFSLMLWVRVDPFTTTRLAGNIIQTCGINC 1080
Db 1021 VIVHLYPFLKGLMGQRNTPPTIVIVWVLLASIFSLMLWVRVDPFTTTRLAGNIIQTCGINC 1080

RESULT 2

US-10-437-963-174885
; Sequence 174885, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(S3221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174885
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72784C.1.pep
; US-10-437-963-174885

Query Match 94.4%; Score 5452.5; DB 16; Length 1081;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 999; Conservative 52; Mismatches 29; Indels 1; Gaps 1;

QY 1 MCGDADALKSGRHGAGDVCOICADGLGTTLDGVFTACDVCRRFPVCRPCYEHKKEGTQA 60
Db 1 MCGDADAVKSGRHSGGQACQICGCGVGTTAEGDVFAACDVCGRPCYCYEYERKDGTOA 60
QY 61 CLQCKTKYKGRHSGPAIRGEGDDTDADDGSDPNYPASGTEDOKQKIADMRMRWMTGG 120
Db 61 CPQCKTKYKGRHSGPAIRGEGDDTDADDVSDYINYPASGSADQKQKIADMRMRWMTAGG 120
QY 121 SGNVGHKPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMGFTGNISRR 180
Db 121 GGDVGRPKYDSGEIGTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMGFTGNIGKR 180
QY 181 APPYVNHSPNPSRFSGSGIGNVANKERVGDGKWKQDKGAIPMTNGTSTIAPSEGRAATDI 240
Db 181 APPYVNHSPNPSRFSGSGIGNVANKERVGDGKWKQDKGAIPMTNGTSTIAPSEGRGVGDI 240
QY 241 DASTEYNMEDALLNDETQPLSRKVPPIASSKNPYRMWIVLRVLVLSIFLHLYRLTNPNVN 300
Db 241 DASTEYNMEDALLNDETQPLSRKVPPIASSKNPYRMWIVLRVLVLSIFLHLYRLTNPNVN 300
QY 301 AYLWLLSVICETWFAWLSWILDQPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
Db 301 AYLWLLSVICETWFAWLSWILDQPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 361 TVDPMKEPPLVTANTVLSILAVDYPVDKVSVCYVSDGAAMLTFDALAETSEFARKWVPFV 420
QY 421 KKYDIEPRAPPEFYFCOKIDYLDKQVOPS FVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 421 KKYNIETPRAPPEWYFSQKIDYLDKXVHPSFVKDRRAMKREYEEFKIRINGLVAKQKVEE 480
QY 481 GWTMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVREKRPFGQHHKKAG 540
Db 481 GWTMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVREKRPFGQHHKKAG 540
QY 541 ANNALVRVSAVLNTQYMLNLDCHYINNSKAVREACFLMDNLPQOVVCYVPPQPFDFG 600
Db 541 ANNALVRVSAVLNTQYMLNLDCHYINNSKALREACFLMDNLPGRSVYVQFPQPFDFG 600
QY 601 IDRNDRVANRNTVFFDINLRGLDGIQGPVYVGTGCVENRTAIYGYEPPIKAKKPG-FLAS 659
Db 601 IDRNDRVANRNTVFFDINLRGLDGIQGPVYVGTGCVENRTAIYGYEPPIKORKKSGFLUSS 660
QY 660 LCGGKKKASKSKRSDDKKSNKHVDSSVPVFNLEDEIEEGVEGAGDFDEKSVLMSQMSLE 719
Db 661 LCGGKKKASKSKRSDDKKSNKHVDSSVPVFNLEDEIEEGVEGAGDFDEKSVLMSQMSLE 720
QY 720 KRFQSAAFVASTLMYEGVGPQSSPTPESSLKKEAIIHVISCGYEDKSEWGTGWIYGSVTE 779
Db 721 KRFQSAAFVASTLMYEGVGPQSSPTPESSLKKEAIIHVISCGYEDKSEWGTGWIYGSVTE 780

```

QY 780 DILTGFMHARGWRSVYCMKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPWY 839
DB 781 DILTGFMHARGWRSVYCMKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPWY 840
QY 840 GYGRUKFLERFAYINTTYPILTSPLLYCYLPAICLTGKFMPEISNLSAIFIALF 899
DB 841 GYGRUKFLERFAYINTTYPILTSPLLYCYLPAICLTGKFMPEISNLSAIFIALF 900
QY 900 LSIFATGILEMRWSGVGIDWWRNEQFWIGGISAHLFAVFOGLLKVLAGIDTNFTVTSK 959
DB 901 LSIFATGILEMRWSGVGIDWWRNEQFWIGGISAHLFAVFOGLLKVLAGIDTNFTVTSK 960
QY 960 ANDEGDFAEALYMKFWTLLIPPTTILINMGVWAGTSYAINSGVQSWGFLGKLPFAF 1019
DB 961 ASDEGDFAEALYMKFWTLLIPPTTILINMGVWAGTSYAINSGVQSWGFLGKLPFAF 1020
QY 1020 WIVHLYPFLKGLMGRQNTPTTIVIVWVALLASIFSLMWVRDPFTRLAGNIOCGIN 1079
DB 1021 WIVHLYPFLKGLMGRQNTPTTIVIVWVALLASIFSLMWVRDPFTRTRVGTQTCGIN 1080
QY 1080 C 1080
DB 1081 C 1081

RESULT 3
US-10-627-132-22
; Sequence 22, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Zea mays
US-10-627-132-22

Query Match 93.9%; Score 5423.5; DB 12; Length 1079;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 996; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

QY 1 MDGDADALKSGRHGAGDYCOICADGIGTLDGVFTACDVCPEPVCRCYBHERKGTQA 60
DB 1 MEGDADGVKSRRGGQVQICGDBGVGTAGDVFACDVCPEPVCRCYBHERKGTQA 60
QY 61 CLOCKTKYKRHSGSPAIRGEGDDTDADDGSDFNYPASGTDQKQIADMRSMWMTGG 120
DB 61 CPQCKNKYKRHSGSPAIRGEGDDTDADDGSDFNYPASGTDQKQIADMRSMWMTGG 120
QY 121 SGNVCHPKYDSGEIGLSKYDSGEIPRGVYVSVTNSQMSGEIPGASPDHMMSPGTNIGR 180
DB 121 SGDVGPRKYDSGEIGLTKYDSGEIPRGVYVSVTNSQMSGEIPGASPDHMMSPGTNIGR 180
QY 181 APFPVNHSPNSREFSGISGNVAKERVGDGKWKQDKGATPMWNTGSIAPSEGRATDI 240
DB 181 APFPVNHSPNSREFSGISGNVAKERVGDGKWKQDKGATPMWNTGSIAPSEGRATDI 240

```

```

QY 241 DASTEYNMEDALINDETQPLSRKVPDIASKINPYRMVIVLRVLVLSIFUHYRLTNVVRN 300
DB 241 DASTDYNMEDALINDETQPLSRKVPDIASKINPYRMVIVLRVLVLSIFUHYRLTNVVRN 300
QY 301 AYPWLWLSVICELWFALSWILDOPFKWFPINRRTYLDRLALRYDREGEPSQLAAVDIFVS 360
DB 301 AYPWLWLSVICELWFALSWILDOPFKWFPINRRTYLDRLALRYDREGEPSQLAAVDIFVS 360
QY 361 TVDPLKEPPIVANTVLSILAVDYPVDKSVSCYVSDGASMLTFDALAETSEFAKRWVPFV 420
DB 361 TVDPMKEPPLVANTVLSILAVDYPVDKSVSCYVSDGASMLTFDALAETSEFAKRWVPFV 420
QY 421 KKYDLEPRAPPEYFCOKIDYLDKQVOPSVFKORRAMKREYEEPKIRINALVSKALVPPE 480
DB 421 KKYDLEPRAPPEYFCOKIDYLDKQVOPSVFKORRAMKREYEEPKIRINALVSKALVPPE 480
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVYREKRPQGHKKAG 540
DB 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVYREKRPQGHKKAG 540
QY 541 ANNALVRVSAVLTNQOYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYQVQPFQFDG 600
DB 541 ANNALVRVSAVLTNQOYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYQVQPFQFDG 600
QY 601 IDNRDRIANRNTVEFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
DB 601 IDNRDRIANRNTVEFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
QY 661 CGGKKKASKSKRSDDKKSKNHVDSSVPVFNLEDEEGVEGAGDPDEKSLVMSQMSLEK 720
DB 661 CGGKKKASKSKK-GSDKKSKQHVDSSVPVFNLEDEEGVEGAGDPDEKSLVMSQMSLEK 719
QY 721 RFGQSAAFVASTLMEYGVVQSSTPESLKEAHIHISCGYEDKSEWGTGIGWYGSVTD 780
DB 721 RFGQSAAFVASTLMEYGVVQSSTPESLKEAHIHISCGYEDKSEWGTGIGWYGSVTD 779
QY 781 ILTGFMHARGWRSVYCMKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPWY 840
DB 781 ILTGFMHARGWRSVYCMKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPWY 839
QY 841 YGGRUKFLERFAYINTTYPILTSPLLYCYLPAICLTGKFMPEISNLSAIFIALF 900
DB 841 YGGRUKFLERFAYINTTYPILTSPLLYCYLPAICLTGKFMPEISNLSAIFIALF 899
QY 901 SIFATGILEMRWSGVGIDWWRNEQFWIGGISAHLFAVFOGLLKVLAGIDTNFTVTSKA 960
DB 901 SIFATGILEMRWSGVGIDWWRNEQFWIGGISAHLFAVFOGLLKVLAGIDTNFTVTSKA 959
QY 961 NDEGDFAEALYMKFWTLLIPPTTILINMGVWAGTSYAINSGVQSWGFLGKLPFAF 1020
DB 961 NDEGDFAEALYMKFWTLLIPPTTILINMGVWAGTSYAINSGVQSWGFLGKLPFAF 1019
QY 1021 VIVHLYPFLKGLMGRQNTPTTIVIVWVALLASIFSLMWVRDPFTRLAGNIOCGIN 1080
DB 1021 VIVHLYPFLKGLMGRQNTPTTIVIVWVALLASIFSLMWVRDPFTRLAGNIOCGIN 1079

RESULT 4
US-10-267-459-6
; Sequence 6, Application US/10267459
; Publication No. US20030150014A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/267,459
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06

```



```
Db 421 KKYIEPRAPWFQKIDYLDKDVHPSFVKDRAMKREVEEFKIRVNGLVAKAQVPEE 480
QY 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSRKRPQGHKKAG 540
Db 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSRKRPQGHKKAG 540
QY 541 AMNALVRVSAVLNNGYMLNLDCHYINNSKAVREAMCFMDNLPQVQYQFPORFDG 600
Db 541 AMNALVRVSAVLNNGYMLNLDCHYINNSKAUREAMCFMDNLPGRSVCYQFPORFDG 600
QY 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPPKAKKPGGLASL 660
Db 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYEPPPKAKKPGGLSSL 660
QY 661 CGGKKKASKSKRSKSKKSNKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEK 720
Db 661 CGGRKKGSKSKK-GSDKKKSQKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEK 719
QY 721 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYGEDKSEWGTGTEIGWYGSVTE 780
Db 720 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYGEDKSEWGTGTEIGWYGSVTE 779
QY 781 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 840
Db 780 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 839
QY 841 YGRLKFLERFAYINTTIYPLTSLPLLYCIIPLAICLLTGKTFMPSINLASIWFIALFL 900
Db 840 YGRLKFLERFAYINTTIYPLTSLPLLYCIIPLAICLLTGKTFMPSINLASIWFISLFI 899
QY 901 SIFATGILEMRWSGVGIDEMWNEQFWLGGI SAHLFAVFOGLLKVLAGIDTNFTVTSKA 960
Db 900 SIFATGILEMRWSGVGIDEMWNEQFWLGGI SAHLFAVFOGLLKVLAGIDTNFTVTSKA 959
QY 961 NDEGDFAELYMFKWTTLLIPPTTILINMVGWVAGTSYAINSGYQSGWGLFGKLFPAFW 1020
Db 960 SDEGDFAELYMFKWTTLLIPPTTILINLVGVAGISYAINSGYQSGWGLFGKLFPAFW 1019
QY 1021 VIVHLYPFLKGLMGRQNRTPPTIVVWALLASIFSLMLWVRDPTTRLAGNIOTCGINC 1080
Db 1020 VIVHLYPFLKGLMGRQNRTPPTIVVWALLASIFSLMLWVRDPTTRLAGNIOTCGINC 1079
```

RESULT 6

```
; Sequence 10, Application US/10627132
; Publication No. US2004006876A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Haiyin
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Zea mays
US-10-627-132-10
```

Query Match

93.8%; Score 5421.5; DB 12; Length 1077;

Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

```
QY 1 MGDADALSKSRHAGDVQCIACDGLTGTLDGDFVFTACDVCRCPCVCEHERKEGTQA 60
Db 1 MEGDADGVKSRRGGGQVQICGDGVTGTTAEGDVFAACDVCGPCVCEYERKGTQA 60
QY 61 CLQCKTKYKRHRGSPAIRGEEGDDTDADGSGFNYPASGTEDQKQIADMRMSRWMTGG 120
Db 61 CPQCKTKYKRHRGSPAIRGEEGDDTDAD--SDFNYLASGNEQKQIADMRMSRWMTGG 118
QY 121 SGNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPCASPDHMHMSTGNSRR 180
Db 119 SGVGRPKYDSGEIGUTKYDSGEIPRGYIPSVTNSQISGEIPCASPDHMHMSTGNIKR 178
QY 181 APFPYVNHSPNPSREFSGSIGNVAMKERVDMKMKODKGAIPTWNTGTSIAPSEGRAATI 240
Db 179 APFPYVNHSPNPSREFSGSIGNVAMKERVDMKMKODKGTIPMTNGTSTAPSEGRGVGI 238
QY 241 DASTYNNMEDALLNDETROPILSRKVPIASSKINPYRMVIVLRVLVLSIFLHYRLTNPVN 300
Db 239 DASTYNNMEDALLNDETROPILSRKVPLFSSRPINPYRMVIVLRVLVLSIFLHYRLTNPVN 298
QY 301 AYPLMLLSVCEIWFALSMLDQFPKWPINRETYLDRALRYDREGEQSOLAAVDIFVS 360
Db 299 AYPLMLLSVCEIWFALSMLDQFPKWPINRETYLDRALRYDREGEQSOLAAVDIFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDYPVKVSCVSDGASMLTFDALAETSEFARKWVPV 420
Db 359 TVDPMKEPPLVANTVLSILAVDYPVKVSCVSDGASMLTFDALAETSEFARKWVPV 418
QY 421 KKYDIEPRAPFPYFQCKIDYLDKDVHPSFVKDRAMKREVEEFKIRVNGLVAKAQVPEE 480
Db 419 KKYDIEPRAPFPYFQCKIDYLDKDVHPSFVKDRAMKREVEEFKIRVNGLVAKAQVPEE 478
QY 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSRKRPQGHKKAG 540
Db 479 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSRKRPQGHKKAG 538
QY 541 AMNALVRVSAVLNNGYMLNLDCHYINNSKAVREAMCFMDNLPQVQYQFPORFDG 600
Db 539 AMNALVRVSAVLNNGYMLNLDCHYINNSKALREAMCFMDNLPGRSVCYQFPORFDG 598
QY 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPPKAKKPGGLASL 660
Db 599 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYEPPPKAKKPGGLSSL 658
QY 661 CGGKKKASKSKRSKSKKSNKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEK 720
Db 659 CGGRKKKASKSKK-GSDKKKSQKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEK 717
QY 721 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYGEDKSEWGTGTEIGWYGSVTE 780
Db 718 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYGEDKSEWGTGTEIGWYGSVTE 777
QY 781 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 840
Db 778 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 837
QY 841 YGRLKFLERFAYINTTIYPLTSLPLLYCIIPLAICLLTGKTFMPSINLASIWFIALFL 900
Db 838 YGRLKFLERFAYINTTIYPLTSLPLLYCIIPLAICLLTGKTFMPSINLASIWFISLFI 897
QY 901 SIFATGILEMRWSGVGIDEMWNEQFWLGGI SAHLFAVFOGLLKVLAGIDTNFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDEMWNEQFWLGGI SAHLFAVFOGLLKVLAGIDTNFTVTSKA 957
QY 961 NDEGDFAELYMFKWTTLLIPPTTILINMVGWVAGTSYAINSGYQSGWGLFGKLFPAFW 1020
Db 958 SDEGDFAELYMFKWTTLLIPPTTILINLVGVAGISYAINSGYQSGWGLFGKLFPAFW 1017
QY 1021 VIVHLYPFLKGLMGRQNRTPPTIVVWALLASIFSLMLWVRDPTTRLAGNIOTCGINC 1080
Db 1021 VIVHLYPFLKGLMGRQNRTPPTIVVWALLASIFSLMLWVRDPTTRLAGNIOTCGINC 1080
```



```
Db 119 SGDVRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMPTGNIGKR 178
QY 181 APPYYNHSPNSRFSGSGIGNVAMKVDGKWKQDKGAIPTWNGTSTAPSEGRAATDI 240
Db 179 APPYYNHSPNSRFSGSGIGNVAMKVDGKWKQDKGTIPTWNGTSTAPSEGRGVGDI 238
QY 241 DASTEYNNMEDALLNDETROPLSRKVP IASKNIPYRMVILVRLVLSIFLHVLRLTNPVN 300
Db 239 DASTDYNMEDALLNDETROPLSRKVP LPSRNIPYRMVILVRLVLSIFLHVLRLTNPVN 298
QY 301 AYPMLLSVICIWFALSWILDOFPKWPFPINRETYLDRALRYDREGESQIAAVIDFVS 360
Db 299 AYPMLLSVICIWFALSWILDOFPKWPFPINRETYLDRALRYDREGESQIAAVIDFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDPVDKVS CVSDGASMLTFDALAETSEFARKWVPV 420
Db 359 TVDPMKPEPLVANTVLSILAVDPVDKVS CVSDGASMLTFDALAETSEFARKWVPV 418
QY 421 KKYDIEPRAPEFYFCOKIDYLDKVP SFVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 419 KKYDIEPRAPEFYFSQIDYLDKVP SFVKDRRAMKREYEEFKVRVNGLVAKAQKVP 478
QY 481 GWIMQDGTFPWGNTRDHPGMIQVFLGHSGLDTEGNELPRLVYVSREKRFQHHKAG 540
Db 479 GWIMQDGTFPWGNTRDHPGMIQVFLGHSGLDTEGNELPRLVYVSREKRFQHHKAG 538
QY 541 ANNALVRYSAVLTNGQYMLNLCDDHYINNSKAVREAMCFMDPNLGPOVCYVQFPQREDG 600
Db 539 ANNALVRYSAVLTNGQYMLNLCDDHYINNSKALREAMCFMDPNLGRSVYVQFPQREDG 598
QY 601 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPIKAKKPGFLASL 660
Db 599 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPIKAKKPGFLSSL 658
QY 661 CGGKKKASKRSDDKKSNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLEK 720
Db 659 CGGRKKASKSKK--GSDKKKSKQKHVDSSVPVFNLEDIEEGVEGAGFDDKSKLLMSQMSLEK 717
QY 721 RFGQSAAFVASTLMYGGVQSGSTPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 780
Db 718 RFGQSAAFVASTLMYGGVQSGSATPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 777
QY 781 ILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVELLSRHCPLWYG 840
Db 778 ILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVELLSRHCPLWYG 837
QY 841 YGGRLKFLERFAYINTTYPILYVTSPLVYLCILPAICLLTGKFIPEISNFASIWFI 900
Db 838 YGGRLKFLERFAYINTTYPILYVTSPLVYLCILPAICLLTGKFIPEISNFASIWFI 897
QY 901 SIFATGILEMRWSGVGIDEWNRNEQFWIGGISAHLFVAVFQGLLAVLAGIDTNFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDEWNRNEQFWIGGISAHLFVAVFQGLLAVLAGIDTNFTVTSKA 957
QY 961 NDEBGDFAELYMKWNTLLIPPTTILINMGVWAGTSYAINSGYQSGWPLFGKLFFAFW 1020
Db 958 SDEGDGFAELYMKWNTLLIPPTTILINMGVWAGTSYAINSGYQSGWPLFGKLFFAFW 1017
QY 1021 VIVHLYPPLKGLMGRQNTPTIIVVAVLLASIFSLLVWVDPPTTRLAGNIPOTCGINC 1080
Db 1018 VIVHLYPPLKGLMGRQNTPTIIVVAVLLASIFSLLVWVDPPTTRVGTGPDOTCGINC 1077
```

RESULT 9

```
US-10-160-719-30
; Sequence 30, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
```

```
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-160-719-30

Query Match 93.8%; Score 5421.5; DB 14; Length 1077;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MGDADALAKSRHGAGDVQCICADGLGTTLDGDVFTACDVCRFPVCRPCYHERKEGTQA 60
Db 1 MEGDADGVKSGRGGQVCQICGDDGVGTGTAEGDVFACDVCVCFVCRPCYERKDGTA 60
QY 61 CLQCKTKYKRHRGSPAIRGEGDDTDADDGSDFNYPASGTEQKOKIADRMESWMTGG 120
Db 61 CQCKTKYKRHRGSPAIRGEGDDTDAD--SDFNYLASGNEQKOKIADRMESWMTGG 118
QY 121 SGNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMPTGNISRR 180
Db 119 SGDVRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMPTGNIGKR 178
QY 181 APPYYNHSPNSRFSGSGIGNVAMKVDGKWKQDKGAIPTWNGTSTAPSEGRAATDI 240
Db 179 APPYYNHSPNSRFSGSGIGNVAMKVDGKWKQDKGTIPTWNGTSTAPSEGRGVGDI 238
QY 241 DASTEYNNMEDALLNDETROPLSRKVP IASKNIPYRMVILVRLVLSIFLHVLRLTNPVN 300
Db 239 DASTDYNMEDALLNDETROPLSRKVP LPSRNIPYRMVILVRLVLSIFLHVLRLTNPVN 298
QY 301 AYPMLLSVICIWFALSWILDOFPKWPFPINRETYLDRALRYDREGESQIAAVIDFVS 360
Db 299 AYPMLLSVICIWFALSWILDOFPKWPFPINRETYLDRALRYDREGESQIAAVIDFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDPVDKVS CVSDGASMLTFDALAETSEFARKWVPV 420
Db 359 TVDPMKPEPLVANTVLSILAVDPVDKVS CVSDGASMLTFDALAETSEFARKWVPV 418
QY 421 KKYDIEPRAPEFYFCOKIDYLDKVP SFVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 419 KKYDIEPRAPEFYFSQIDYLDKVP SFVKDRRAMKREYEEFKVRVNGLVAKAQKVP 478
QY 481 GWIMQDGTFPWGNTRDHPGMIQVFLGHSGLDTEGNELPRLVYVSREKRFQHHKAG 540
Db 479 GWIMQDGTFPWGNTRDHPGMIQVFLGHSGLDTEGNELPRLVYVSREKRFQHHKAG 538
QY 541 ANNALVRYSAVLTNGQYMLNLCDDHYINNSKAVREAMCFMDPNLGPOVCYVQFPQREDG 600
Db 539 ANNALVRYSAVLTNGQYMLNLCDDHYINNSKALREAMCFMDPNLGRSVYVQFPQREDG 598
QY 601 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPIKAKKPGFLASL 660
Db 599 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPIKOKKGGFLSSL 658
QY 661 CGGKKKASKRSDDKKSNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLEK 720
Db 659 CGGRKKASKSKK--GSDKKKSKQKHVDSSVPVFNLEDIEEGVEGAGFDDKSKLLMSQMSLEK 717
QY 721 RFGQSAAFVASTLMYGGVQSGSTPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 780
Db 718 RFGQSAAFVASTLMYGGVQSGSATPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 777
QY 781 ILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVELLSRHCPLWYG 840
```

Db 778 ILTGKMHARGNSIYCMKPAKGSAPINLSDRNLQVRLWALGSEVILFSRHCPLWYG 837
Qy 841 YGRLKFLERFAYINTIYPLTSLPLLVYLCILPAICLLTGKFMPIBSINLASIWFIALFL 900
Db 838 YGRLKFLERFAYINTIYPLTSLPLLVYLCILPAICLLTGKFMPIBSINLASIWFIALFL 897
Qy 901 SIFATGILEMRWSGVGIDSEWRNEQFVWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDSEWRNEQFVWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA 957
Qy 961 NDEEGDFAELYMFKWTTLLIPPTTILINMVGVSAGTSYAINSGYOSWGFLFKGLFFAFW 1020
Db 958 SDEGDGFAELYMFKWTTLLIPPTTILINLVGVVAGISYAINSGYOSWGFLFKGLFFAFW 1017
Qy 1021 VIVHLYPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRIDPFTTRVTGPDQTGGINC 1080
Db 1018 VIVHLYPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRIDPFTTRVTGPDQTGGINC 1077

RESULT 10
US-10-160-719-50
; Sequence 50, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-50

Query Match 91.1%; Score 5261.5; DB 14; Length 1043;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 970; Conservative 42; Mismatches 31; Indels 3; Gaps 2;

Qy 1 MCGDADALSGRGAGDVCOICADGLGTTLDGDFVFTACDVCRFPVCRPCYEHERKEGTQA 60
Db 1 MEGDAGVSGRGGGQVCOICDGVGTTTAEAGDVFAACDVCGFPVCRPCYEHERKEGTQA 60
Qy 61 CLQCKTKYKRRHSGPAIRGEGDDTDADGSDGFNYPASGTEDQKQIADRRMRWRMTGG 120
Db 61 CPQCKTKYKRRHSGPAIRGEGDDTDAD--SDFNYLASGNEQKQIADRRMRWRMTGG 118
Qy 121 SGNVGHPKYSIGILSKYDSGELPRGYVPSVTNSQMSGIPGASPDHNMWSPGTNLSRR 180
Db 119 SGDVGRPKYDSIGILTKYDSGELPRGYVPSVTNSQMSGIPGASPDHNMWSPGTNIGKR 178
Qy 181 APPYVNHSPNPREFSGSIGVNAWKERVGVKMKQDKGAIPTWNGTSIAPSEGRATDI 240
Db 179 APPYVNHSPNPREFSGSIGVNAWKERVGVKMKQDKGAIPTWNGTSIAPSEGRGVCDI 238
Qy 241 DASTEYNNEDALLNDETQPLSRKVPFIASSKINPYRMVILVRLVLVLSIFLHYRLTNPVN 300
Db 239 DASTDYNNEDALLNDETQPLSRKVPFIASSKINPYRMVILVRLVLVLSIFLHYRLTNPVN 298
Qy 301 AYLWLLSVCEIWFALSWILDQPKWFFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
Db 299 AYLWLLSVCEIWFALSWILDQPKWFFPINRETYLDRALRYDREGEPSQLAAVDIFVS 358

Qy 361 TVDPLKEPPIVANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 359 TVDPMKEPPLVANTVLSILAVDYPVDKVSVCYVSDDGAAMLTFDALAETSEFARKWVPFV 418
Qy 421 KKYDIEPRAPPEYFOKIDYLDKQVQSPVKDORRAMKREYEBKIRINALVSKALKVPEE 480
Db 419 KKYNIERAPPEYFOSKIDYLDKQVHPSFVKDORRAMKREYEEFKVNVGLVAKAQVPEE 478
Qy 481 GWMODGTWPCNTRDHPGMLQVFLGHSGLDTEGNELPRLVVYVSRKRPQFQHHKAG 540
Db 479 GWMQDGTWPCNTRDHPGMLQVFLGHSGLDTEGNELPRLVVYVSRKRPQFQHHKAG 538
Qy 541 ANMALVRVSAVLITNGQYMLNLDCHYINNSKAVREAMCFMLDNLGPEQVCYVQPPORFDG 600
Db 539 ANMALVRVSAVLITNGQYMLNLDCHYINNSKALREAMCFMLDNLGSRVCYVQPPORFDG 598
Qy 601 IDNRDRIYANRVVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPIKAKKPEGLASL 660
Db 599 IDNRDRIYANRVVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPIKAKKPEGLSS 658
Qy 661 CGGKKKASKSKSSDKKSNKHVDSSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEK 720
Db 659 CGGKKKASKSKK--GSDKKSKOKHVDSSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEK 717
Qy 721 RFGQSAAFVASTILMEYGGVPOQSATPESLLKEAIIHVISCGYEDKTEWGTGTEIGMIVGSVTD 780
Db 718 RFGQSAAFVASTILMEYGGVPOQSATPESLLKEAIIHVISCGYEDKTEWGTGTEIGMIVGSVTD 777
Qy 781 ILTGFKHARGHRSVYCMKPAKGSAPINLSDRNLQVRLWALGSEVILFSRHCPLWYG 840
Db 778 ILTGFKHARGHRSYIYCMKPAKGSAPINLSDRNLQVRLWALGSEVILFSRHCPLWYG 837
Qy 841 YGRLKFLERFAYINTIYPLTSLPLLVYLCILPAICLLTGKFMPIBSINLASIWFIALFL 900
Db 838 YGRLKFLERFAYINTIYPLTSLPLLVYLCILPAICLLTGKFMPIBSINLASIWFIALFL 897
Qy 901 SIFATGILEMRWSGVGIDSEWRNEQFVWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDSEWRNEQFVWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA 957
Qy 961 NDEEGDFAELYMFKWTTLLIPPTTILINMVGVSAGTSYAINSGYOSWGFLFKGLFFAFW 1020
Db 958 SDEGDGFAELYMFKWTTLLIPPTTILINLVGVVAGISYAINSGYOSWGFLFKGLFFAFW 1017
Qy 1021 VIVHLYPFLKGLMGRQNRPTIIVVW 1046
Db 1018 VIVHLYPFLKGLMGRQNRPTIIVVW 1043

RESULT 11
US-10-627-132-14
; Sequence 14, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14

```
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Zea mays
US-10-627-132-14

Query Match
Best Local Similarity 89.1%; Score 5149; DB 12; Length 1076;
Matches 948; Conservative 69; Mismatches 58; Indels 6; Gaps 6;

QY 1 MDGDADALKSRHGAGDVCCQCADGLGTTLDGVDFTACDVCRFPVCRPCYEHHERKEGTQA 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MDG-GDATNSGKHVAGQVCQICGDGVGTAAADGLFTACDVCGFPVCRPCYEHKDGTA 59

QY 61 CLQCKTKYKRHRGSPAIRGEEGDDTDADDGSDFNYPASGTEDQKQIADRMRSWRMTGG 120
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
60 CPQCKTKYKRHKSGPPVHGEENEDVDADDVSDYNYQASGNQDQKQIAERMLTWTNREG 119

QY 121 SGNVGHKYDSGEIGLSKYDSGEIPRGYVPSVNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
120 S-DIGLAKYDSGEIGHGKYDSGEIPRGYIPSLTHSQISGEIPGASPD-HMMSVPGNIGRR 177

QY 181 A-PPFYVNHSPNPSREFSGSIGNVAKKVDGWKMKQDKGAIPMTNGTISIAPSEGRAATD 239
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
178 GHQFFYVNHSPNPSREFSGSLGNVAKKVDGWKMK-DKGAIPMTNGTISIAPSEGRVAD 236

QY 240 IDASTEYNMEDALLNDETROPQLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR 299
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
237 IDASTDYNMEDALLNDETROPQLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR 296

QY 300 NAYPLWLLSVICEIWFALSWILDQFPKWFPIINRETYLDRALRYDREGEPSSQLAAVDIFV 359
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
297 NAYPLWLLSVICEIWFALSWILDQFPKWFPIINRETYLDRALRYDREGEPSSQLAPVDIFV 356

QY 360 STVDPLKEPPIVANTVLSILAVDPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF 419
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
357 STVDPMKEPPIVANTVLSILAVDPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF 416

QY 420 VKKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALKVPE 479
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Query Match 89.1%; Score 5149; DB 14; Length 1076;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 948; Conservative 69; Mismatches 58; Indels 6; Gaps 6;

QY 1 MDGDADALKSRHGAGDVCCQCADGLGTTLDGVDFTACDVCRFPVCRPCYEHHERKEGTQA 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MDG-GDATNSGKHVAGQVCQICGDGVGTAAADGLFTACDVCGFPVCRPCYEHKDGTA 59

QY 61 CLQCKTKYKRHRGSPAIRGEEGDDTDADDGSDFNYPASGTEDQKQIADRMRSWRMTGG 120
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
60 CPQCKTKYKRHKSGPPVHGEENEDVDADDVSDYNYQASGNQDQKQIAERMLTWTNREG 119

QY 121 SGNVGHKYDSGEIGLSKYDSGEIPRGYVPSVNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
120 S-DIGLAKYDSGEIGHGKYDSGEIPRGYIPSLTHSQISGEIPGASPD-HMMSVPGNIGRR 177

QY 181 A-PPFYVNHSPNPSREFSGSIGNVAKKVDGWKMKQDKGAIPMTNGTISIAPSEGRAATD 239
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
178 GHQFFYVNHSPNPSREFSGSLGNVAKKVDGWKMK-DKGAIPMTNGTISIAPSEGRVAD 236

QY 240 IDASTEYNMEDALLNDETROPQLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR 299
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
237 IDASTDYNMEDALLNDETROPQLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR 296

QY 300 NAYPLWLLSVICEIWFALSWILDQFPKWFPIINRETYLDRALRYDREGEPSSQLAAVDIFV 359
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
297 NAYPLWLLSVICEIWFALSWILDQFPKWFPIINRETYLDRALRYDREGEPSSQLAPVDIFV 356

QY 360 STVDPLKEPPIVANTVLSILAVDPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF 419
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
357 STVDPMKEPPIVANTVLSILAVDPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF 416

QY 420 VKKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALKVPE 479
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
417 CKYNIIEPRAPEFYFAQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALKVPE 476

QY 480 EGWIMQDTPMGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSREKRPFGFHKKKA 539
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
477 EGWIMQDTPMGNTRDHPGMIQVFLGHSGGLDVEGNEPLRLVYVSREKRPFGFHKKKA 536

QY 540 GAMNALVRVSAVLITNGQYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYVQPPQF 599
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
537 GAMNALVRVSAVLITNGQYMLNLDCHYINNSKALREAMCFMDPNLGRNVYVQPPQF 596

QY 600 GIDRNDRYANRNTVFFDNLRLGLDGIQGPVYVGTGCVFNRTAIYGEPPIKAKKPGFLAS 659
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
597 GIDRNDRYANRNTVFFDNLRLGLDGIQGPVYVGTGCVFNRTAIYGEPPVKKKPGFFSS 656

QY 660 LCGGKKKASKKRSDDKKKNKHVDSSVPFNLEDIEEGVEGAGFDDEKSVLMSQMSLE 719
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
657 LCGGRKTSKSKK-SSEKKSHRHADSSVPFNLEDIEEGIEGSGQFDDKSLMSQMSLE 715

QY 720 KRFGSAAFAVSTLMYEGVPOSSTPESLLKBAIHVISCYEDKSEWGTETGIWYGSVTE 779
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
716 KRFGSSVFVSTLMYEGVPOSATPESLLKBAIHVISCYEDKTDWGTETGIWYGSVTE 775

QY 780 DILTGFKHARGWRSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSPRHCPLWY 839
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
776 DILTGFKHARGWRSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSPRHCPIWY 835

QY 840 GYGRKLKFLERFAYINTTIYPLTSLFLVYICILPAICLLTGKFINMPEISNLASIFWIFALF 899
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
836 GYGRKLKFLERFAYINTTIYPLTSLFLVYICILPAICLLTGKFIIPKISNLESVWFISLF 895

QY 900 LSI FATGILEMRWSVGDEWNRNQFVWIGISIAHLFAVFOGLLKVLVLAGIDTNTVTTSK 959
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
896 LSI FATGILEMRWSVGDEWNRNQFVWIGISIAHLFAVFOGLLKVLVLAGIDTNTVTTSK 955
```

Db 417 CKKYNIEPRAPEWYFAQKIDYLDKQVOTSFVKERRAMKREYEBFKVRINGLVAKAQKVE 476
QY 480 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKSPGFQHHKA 539
Db 477 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDVEGNEPLRLVYVSREKSPGFQHHKA 536
QY 540 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMIDPNLGPOVCYVQFPQFED 599
Db 537 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMIDPNLGRNVYVQFPQFED 596
QY 600 GIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPKAKKPGFLAS 659
Db 597 GIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPVKKKKPGFSS 656
QY 660 LCGGKKKSKSKKSSDKKSNKHVDSSVPFNLEDEIEGIEGSDDEKSLIMSOMSL 719
Db 657 LCGGKKKSKSKK-SSEKKKSHRHADSSVPFNLEDEIEGIEGSDDEKSLIMSOMSL 715
QY 720 KREGQSAAFVASTLMEYGGVPOQSTPESLLKEAHHVISCYEDKSEWGTGIEGMYGSVTE 779
Db 716 KREGQSSVFVASTLMEYGGVPOQSATPESLLKEAHHVISCYEDKTDWGTGIEGMYGSVTE 775
QY 780 DILTFKMHARGWRSYVCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCPFIW 839
Db 776 DILTFKMHARGWRSYVCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCPFIW 835
QY 840 GYGRLKFLERFAYINTIYPLTSLPLVYCIIPALCLLTGKFIEMPEISNLASIMFIALF 899
Db 836 GYGRLKFLERFAYINTIYPLTSLPLVYCIIPALCLLTGKFIEMPEISNLASIMFIALF 895
QY 900 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLEFAVFGQLLKVLAGIDTFTVTSK 959
Db 896 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLEFAVFGQLLKVLAGIDTFTVTSK 955
QY 960 ANDEEGDFAELYNFKMTLLIPIPTTILINMVGWAGTSYAINSGYQSWGFLFKLFFAF 1019
Db 956 ATDEEGDFAELYNFKMTLLIPIPTTILINMVGWAGTSYAINSGYQSWGFLFKLFFAF 1015
QY 1020 WVIVHLYPFLKGLMGKQNRPTTIVVWALLASIFSLMLWRVDPFTTRLAGPNIQTCGIN 1079
Db 1016 WVIVHLYPFLKGLMGKQNRPTTIVVWALLASIFSLMLWRVDPFTTRLAGPNIQTCGIN 1075
QY 1080 C 1080
Db 1076 C 1076

RESULT 13

US-10-160-719-58
; Sequence 58, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-58

Query Match 89.1%; Score 5149; DB 14; Length 1076;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 948; Conservative 69; Mismatches 58; Indels 6; Gaps 6;
QY 1 MDGADALXKGRHAGDVQCI CADGLGTLTLDGVDFTACDVCRFPVCRPCYEHEHKEGTQA 60
Db 1 MDG-GDATNSGKHVAGVCQICGDBGVTAADGDLFTACDVCGPVCPCYEHEHKEGTQA 59
QY 61 CLOCKTKYRHRHSGPAIRGBEGDDTDADGSDPNYPASGTEDOKOIAIDMRMRWMTGG 120
Db 60 CPOCKTKYRHRHSGSPVHGEENEDVDDSDNYQASGNQOKOIAERMLTWRNRSRG 119
QY 121 SGNVGHKPYDSGEIGHLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPTGISRR 180
Db 120 S-DIGLAKYDSGEIGHGKYDSGEIPRGYIPLSTHSGIISGEIPGASPD-HMMSVPGNIGRR 177
QY 181 A-PPFYVNSHPNPSRFSGSI GNANKERVDGKMKQDKGAIPTMTNGTSTAPSEGRAATD 239
Db 178 GHOPFYVNSHPNPSRFSGSLGNVANKERVDGKMK-KDGAIPTMTNGTSTAPSEGRGVAD 236
QY 240 IDASTEYNMEDALLNDETROPISRKVPKPIASSKINPYRMVILRLVLSIFLHYRLTNPVR 299
Db 237 IDASTEYNMEDALLNDETROPISRKVPKPISSRINPYRMVILRLVLSIFLHYRLTNPVR 296
QY 300 NAYPLMLLSVCEIWFALSMILDOFPKWPFPINRETYLDRALRYDREGESQIAAADIVF 359
Db 297 NAYPLMLLSVCEIWFALSMILDOFPKWPFPINRETYLDRALRYDREGESQIAAADIVF 356
QY 360 STVDPLKEPPIVTANTVLSILAVDYPDKVSCYVSDGASMLTFDALAETSEFARKWVPF 419
Db 357 STVDPMKEPPLVTANTVLSILAVDYPDKVSCYVSDGASMLTFDALAETSEFARKWVPF 416
QY 420 VKKYDIEPRAPEYFCOKJDYLDKQVOPSFVKDRAMKREYEBFKIRINALVSKALKVPE 479
Db 417 CKKYNIEPRAPEWYFAQKIDYLDKQVOTSFVKERRAMKREYEBFKVRINGLVAKAQKVE 476
QY 480 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKSPGFQHHKA 539
Db 477 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDVEGNEPLRLVYVSREKSPGFQHHKA 536
QY 540 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMIDPNLGPOVCYVQFPQFED 599
Db 537 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMIDPNLGRNVYVQFPQFED 596
QY 600 GIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPKAKKPGFLAS 659
Db 597 GIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPVKKKKPGFSS 656
QY 660 LCGGKKKSKSKKSSDKKSNKHVDSSVPFNLEDEIEGIEGSDDEKSLIMSOMSL 719
Db 657 LCGGKKKSKSKK-SSEKKKSHRHADSSVPFNLEDEIEGIEGSDDEKSLIMSOMSL 715
QY 720 KREGQSAAFVASTLMEYGGVPOQSTPESLLKEAHHVISCYEDKSEWGTGIEGMYGSVTE 779
Db 716 KREGQSSVFVASTLMEYGGVPOQSATPESLLKEAHHVISCYEDKTDWGTGIEGMYGSVTE 775
QY 780 DILTFKMHARGWRSYVCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCPFIW 839
Db 776 DILTFKMHARGWRSYVCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCPFIW 835
QY 840 GYGRLKFLERFAYINTIYPLTSLPLVYCIIPALCLLTGKFIEMPEISNLASIMFIALF 899
Db 836 GYGRLKFLERFAYINTIYPLTSLPLVYCIIPALCLLTGKFIEMPEISNLASIMFIALF 895
QY 900 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLEFAVFGQLLKVLAGIDTFTVTSK 959
Db 896 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLEFAVFGQLLKVLAGIDTFTVTSK 955
QY 960 ANDEEGDFAELYNFKMTLLIPIPTTILINMVGWAGTSYAINSGYQSWGFLFKLFFAF 1019
Db 956 ATDEEGDFAELYNFKMTLLIPIPTTILINMVGWAGTSYAINSGYQSWGFLFKLFFAF 1015
QY 1020 WVIVHLYPFLKGLMGKQNRPTTIVVWALLASIFSLMLWRVDPFTTRLAGPNIQTCGIN 1079

Db 1016 WVIVHLYPFLKGLMGKQNRTPTIIVVWAILLASIFSLMVRIDPFTTRVTGPDIAKCGIN 1075
QY 1080 C 1080
Db 1076 C 1076

RESULT 14

US-10-425-114-58605
; Sequence 58605, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; LENGTH: 1119
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700217152_FLI.pep
US-10-425-114-58605

Query Match 88.8%; Score 5129; DB 12; Length 1119;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 944; Conservative 71; Mismatches 60; Indels 6; Gaps 6;

QY 1 MDGADALSKGRHAGDVCOICADGLGTTLDGDVFTACDVCRFPVCRPCYHERKEGTQA 60
Db 44 MDG-GDAYSNGKHVAGQVCICGQGVGTAAADGLFTACDVCRFPVCRPCYHERKGTQA 102
QY 61 CLQCKTKYKRRHSGPAIRGEEDDDTADGSDFNYPASGTDQKQKADRMESMNMNTGG 120
Db 103 CPQCKTKYKRRHSGPVPVHEENEDVDADDVSDYNSQASGQDQKQKIAERMLTWITNSRG 162
QY 121 SGNVCHPKYDSGEICLSKYDSGEIPRGVPSVNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db 163 S-DIGLAKYDSGEIGHGYDSGEIPRGVPSLTHSQISGEIPGASPD-HMMSPVGNIGGR 220
QY 181 A-PPYPVNHSPNPSREFSGSIGNVAKERVGVGKMKQDKGAIPMTNGTSIAPSEGRAATD 239
Db 221 GHQFPYVNHSPNPSREFSGSIGNVAKERVGVGKMK-DKGALPMTNGTSIAPSEGRGVAD 279
QY 240 IDASTEYNMEDALLNDETQPLSRKVPPIASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR 299
Db 280 IDASTDYMEDALLNDETQPLSRKVPPISSRINPYRMVIVLRVLVLCIFLRYRITHFVN 339
QY 300 NAYPLWLLSVICEWFWALSILWLDQPKWFPINRETYLDRLALRYDREGEPQLAAVDIFV 359
Db 340 NAYPLWLLSVICEWFWALSILWLDQPKWSPINRETYLDRLALRYDREGEPQLAPVDIFV 399
QY 360 STVDPLKEPPIVTANTVLSILAVDYPVKVCYVSDGASMLTDLALETSEFARKVVPF 419
Db 400 STVDPMKEPPIVTANTVLSILAVDYPVKVCYVSDGGAAMLTDLALETSEFARKVVPF 459
QY 420 VKKYDIEPRAPEFYCQKIDYLDKQVPSFVKDRRAMKREYEEFKIRINALVSKALKVPE 479
Db 460 CKKYNIEPRAPEWYFAQIDYLDKQVTSFVKERRAMKREYEEFKVRLINGLVAQAQVPE 519
QY 480 EGMWQDGTTPGNNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVSREKRPQGHKKKA 539
Db 520 EGMWQDGTTPGNNTRDHPGMIQVFLHSGGLDVEGNEPRLVYVSREKRPQGHKKKA 579

QY 540 GAMNALVRYSAVLTNGQYMLNLDCHYINNKAVERAMCFMLDPNLPQVCYVQFPQRD 599
Db 580 GAMNALVRYSAVLTNGQYMLNLDCHYINNKAVERAMCFMLDPNLPGRVNCYVQFPQRD 639
QY 600 GIDRNDRYANRNTVPFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLAS 659
Db 640 GIDRNDRYANRNTVPFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPVKKKPGFFS 699
QY 660 LCGGKKKASKKRSDDSKKSNKVDSSVPFNLEDIEEGVEGAGFDDKSVLMSQMSLE 719
Db 700 LCGRKTTSKSKK-SSEKKESHADSSVPFNLEDIEEGSQFDDKSLMSQMSLE 758
QY 720 KEPGQSAFVASTLMYEGVPOSSTPESLLKEAIIHVISGYEDKSWGTEIGWIYGSVTE 779
Db 759 KRFQSSVFVASTLMYEGVPOSATPESLLKEAIIHVISGYEDKTDWGTGTEIGWYGSVTE 818
QY 780 DILTFGKHARGWRSVYCMKRPAPKGSAPINLSRDLNOLRWALGSEILFSRHCPWIY 839
Db 819 DILTFGKHARGWRSVYCMKRPAPKGSAPINLSRDLNOLRWALGSEILFSRHCPWIY 878
QY 840 GYGBRLKFLERFAYINTTIYPLTSLPLVYCILPAICLLTGKFIIMPEISNLASIWFIALF 899
Db 879 GYGBRLKFLERFAYINTTIYPLTSLPLVYCILPAICLLTGKFIIMPEISNLASIWFIALF 938
QY 900 LSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLPFAVFOGLLKVLADITNFTVTSK 959
Db 939 LSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLPFAVFOGLLKVLADITNFTVTSK 998
QY 960 ANDERGEDFAELYMFKWTTLLIPPTTILLINMGVVAGTSYAINSGVQSGPLFGKLFPAF 1019
Db 999 ATDESGDFAELYMFKWTTLLIPPTTILLINMGVVAGTSYAINSGVQSGPLFGKLFPAF 1058
QY 1020 WVIVHLYPFLKGLMGKQNRTPTIIVVWAILLASIFSLMVRIDPFTTRVTGPDIAKCGIN 1079
Db 1059 WVIVHLYPFLKGLMGKQNRTPTIIVVWAILLASIFSLMVRIDPFTTRVTGPDIAKCGIN 1118
QY 1080 C 1080
Db 1119 C 1119

RESULT 15

US-10-437-963-142245
; Sequence 142245, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142245
; LENGTH: 1063
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1063)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4326C.1.pep
US-10-437-963-142245

Query Match 88.7%; Score 5122.5; DB 16; Length 1063;

Best Local Similarity 87.9%; Pred. No. 0;									
Matches 943; Conservative 64; Mismatches 51; Indels 15; Gaps 4;									
QY	9	KSGRHGAGVQICADGLGTTLDGVFTACDVCRCPCYHERKEGTQACLOCKTKY	68						
Db	5	KSGRHGSGQACQILGDGX-----DVGFPVCRPCYERKDGSOACPCQCKTKY	52						
QY	69	KHRGSPAIRGEGDDTDADGDSDNYFASGTEDQKTIADRMRSWRMNTGGSGNVGHPK	128						
Db	53	KRHGSPPIILGDESDDVDDADDASDVNYPTSGQDQHKHKAERMLTWRMNSGRDDIVHVK	112						
QY	129	YDSGEIGLSKSDSGEIPRGVPSVINSQMSGEIPGASPDHMHMSPTGNISSRA-PPPYVN	187						
Db	113	YDSGEIGHPKYDSGEHPIRYIPLSLTHSQISGEIPGASPD-HMMSPVGNIGRRGHPPPYVN	171						
QY	188	HSPNPSREFSGSIGNVAKERVQWQKQKGAIPMTNGTSTAPSEGRAATDIDASTEYN	247						
Db	172	HSPNPSREFSGSIGNVAKERVQWQK-KGAIPMANGTSTAPSEGRGVGDIDASTDYN	230						
QY	248	MEDALINDETROPLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVRNAYPLWLL	307						
Db	231	MEDALINDETROPLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVRNAYPLWLL	290						
QY	308	SVICEIWFALSWILDQFPKWFPIINRETYLDRLALRYDREGEPQSLAAVDIFVSTVDPLKE	367						
Db	291	SVICEIWFALSWILDQFPKWSPINRETYLDRLALRYDREGEPQSLAPVDIFVSTVDPKME	350						
QY	368	PPIVTANTVLSILAVDYPVDKVCYVSDDGASMLTFDALAETSEFARKVPPVVKYDIEP	427						
Db	351	PELVANTANTVLSILAVDYPVDKVCYVSDDGAAMLTFDALAETSEFARKVPPVCKKYSIEP	410						
QY	428	RAPEFYFCOKIDYLDKQVOPSFKVDRRAMKREYEEFKIRINALVSKALKVPEEGWIMQDG	487						
Db	411	RAPEWYFAQIDYLDKQVQASFKVDRRAMKREYEEFKVRNALVAKAQKYPEEGWIMQDG	470						
QY	488	TPWPGNNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVSREKRPFGOHKKKAGAMNALVR	547						
Db	471	TPWPGNNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVSREKRPFGOHKKKAGAMNALVR	530						
QY	548	VSALVTNGQYMLNDCDHYINNSKAVREACFLMDPNLGPQVCYVQPPQDFGDIRNDRY	607						
Db	531	VSALVTNGQYLLNDCDHYINNSKALREACFLMDPNLGRRCYVQPPQDFGDIRNDRY	590						
QY	608	ANRNTVFFDINLRGLDGIQGVVGTGCVFNRTAIYGYEPPYIKAKKPGFLASLCGGKKKA	667						
Db	591	ANRNTVFFDINLRGLDGLQGVVGTGCVFNRTALYGYEPPYIKQKRPGYFSSLCGGKKKT	650						
QY	668	SKSKRSKSSKKSNKHVDSSVPVFNLEDEEGVEGAGFDEKSVLMSQMSLEKRFQOSAA	727						
Db	651	KKSKEKSTKKSHKHVDSSVPVFNLEDEEGIEGSGFDDDEKSLMSQMSLEKRFQOSSV	710						
QY	728	FVASTLMYGGVQOSSTPESLLKEAITHVISCYGEDKSEWTEIGWYGVSTEDILTCFKM	787						
Db	711	FVASTLMYGGVQOSATPESLLKEAITHVISCYGEDKSDWTEIGWYGVSTEDILTCFKM	770						
QY	788	HARGWRSVCMKRPAPKPGSAPINLSRLNQVLRWALGVSVEILFSRHCPLWYGYGRLKF	847						
Db	771	HARGWRSIYCMKRPAPKPGSAPINLSRLNQVLRWALGVSVEILFSRHCPWYGYGRLKF	830						
QY	848	LERFAYINTTIYPLTSLPLVYLCILPAICLLTGKFMPEISNLSIASIWFIALFISIFATGI	907						
Db	831	LERFAYINTTIYPLTSLPLLYLCILPAICLLTGKFIIPESISNFSIASIWFISLISIFATGI	890						
QY	908	LEMWWSGVGDDEWRNEQFVWIGISAHLEFAVFOGLKVLGADITNETVTTSKANDERGDF	967						
Db	891	LEMWWSGVGDDEWRNEQFVWIGISAHLEFAVFOGLKVLGADITFTVTSKASDEERGDF	950						
QY	968	AELYMFKWTTLLIPTTILLINMVGVVAGTSYAINSGVQSGMPLFGKLFPAFWIVHLYP	1027						
Db	951	AELYMFKWTTLLIPTTILLINMVGVVAGTSYAINSGVQSGMPLFGKLFPAFWIVHLYP	1010						
QY	1028	FLKGLMGQRNPTTIVVAVLLASIFSLMWVRVDPFTTRLAGNIOTCGINC	1080						

Db 1011 FLKGLMGQRNPTTIVVAVLLASIFSLMWVRVDPFTTRVGPDTQKCGINC 1063

Search completed: August 23, 2004, 01:17:23
Job time : 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 15:49:02 ; Search time 9999 Seconds
(without alignments)
4681.516 Million cell updates/sec

Title: US-09-900-237A-30
Perfect score: 5778
Sequence: 1 MDGDADALKSGRHGADVQC.....VDPFTTRLAGPNIQTCGINC 1080

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09900237/runat_18082004_081514_14633/app_query.fasta_l.1223
-DB=genEmbl -QMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -IOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09900237 @CGN_1_1_6626@runat_18082004_081514_14633 -NCFU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: go_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rod.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5778	100.0	3626	8	BT009438
2	5452.5	94.4	3954	8	AK072356
3	5434.5	94.1	3264	6	AX653232
4	5423.5	93.9	3795	8	AF200533
5	5423.5	93.9	3799	6	AX338680
6	5421.5	93.8	3745	8	AF200528
7	5169.5	89.5	3222	6	AX652952
8	5165.5	89.4	4282	8	AK069196
9	5143	89.0	3676	8	AF200529
10	4838	83.7	87792	8	AP003837
11	4632	80.2	41693	2	AC145384
12	4632	80.2	154555	8	AC135958
13	4577.5	79.2	3723	8	AF150630
14	4530	78.4	3229	8	BT002335
15	4507	78.0	3682	8	AF027174
16	4499	77.9	3614	6	AR267559
17	4499	77.9	3614	6	AX030946
18	4499	77.9	3614	6	BD022678
19	4458	77.2	3532	8	AY055724
20	4089.5	70.8	29292	8	AB018111
21	4008	69.4	3640	8	AY162181
22	3993.5	69.1	3725	8	AF200526
23	3992.5	69.1	3732	8	AK099228
24	3992.5	69.1	3764	8	AK098978
25	3992.5	69.1	3768	8	AK102140
26	3992.5	69.1	3802	8	AK067967
27	3989.5	69.0	3801	8	AK099281
28	3981.5	68.9	3897	8	AK100188
29	3981.5	68.9	3752	8	AF200525
30	3963.5	68.6	3603	6	AR267557
31	3963.5	68.6	3603	6	AX030942
32	3963.5	68.6	3603	6	BD022676
33	3963.5	68.6	3763	8	BT008654
34	3959.5	68.5	3673	6	AR267560
35	3959.5	68.5	3673	6	AX030948
36	3959.5	68.5	3673	6	BD022679
37	3946	68.3	3851	6	BD236020
38	3870.5	67.0	3631	8	AK121170
39	3844.5	66.5	3812	8	AF200532
40	3840.5	66.5	4127	8	AK121193
41	3828	66.3	4029	8	AK100877
42	3823	66.2	3746	8	AK100914
43	3820	66.1	3538	8	AF200530
44	3818	66.1	3968	8	AF200531
45	3815.5	66.0	4208	8	AK073561

ALIGNMENTS

BT009438
 LOCUS 3626 bp mRNA linear PLN 20-JUN-2003
 DEFINITION Triticum aestivum clone wimk4.pk0015.all:fis, full insert mRNA sequence.
 ACCESSION BT009438
 VERSION BT009438.1 GI:32128989
 KEYWORDS FLI CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 3626)
 Tingley,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H., Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
 AUTHORS Direct Submission
 TITLE Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 JOURNAL USA

FEATURES
 source Location/Qualifiers
 1..3626
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wimk4.pk0015.all:fis"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3626
 Score: 5778.00 Matches: 1080
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0

US-09-900-237A-30 (1-1080) x BT009438 (1-3626)

QY	1	MetAspGlyAspAlaAspAlaLeuLysSerGlyValArgHisGlyAlaGlyAspValCysGln	20
Db	60	ATGACCGCGGACCGCGGACGCTCGAAGTCCGGGAGGACCGGGCCGGGACGTGTGCAG	119
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
Db	120	ATCTGCCGCGACGCGCTGGGACACACAGTGTGGACGCGAGCTTCAACGCGCTGCGAGCTC	179
QY	41	CysArgPheProValCysArgProCysTyrClnHisGluArgLysGluGlyThrGlnAla	60
Db	180	TGCGCGCTTCCGCGCTGCGCGCCCTGCTACGAGCACGAGGCGCAAGGAGGCGCCAGGCC	239
QY	61	CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
Db	240	TGCTCCAGTCCAGACCAAGTACAGCGCCACAGAGGAGCCAGCGATCCGCGGGGAG	299
QY	81	GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
Db	300	GAAGCGCAGCAGACTGATGCGCGATGATGTAGTGACTTCAACTACCCCTGCATCTGGCACT	359
QY	101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly	120
Db	360	GAGGACCAAGAGCAGAGGATGCTGACAGGATGCGCAGCTGGCGCATGAACACCGGGGGC	419
QY	121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
Db	420	AGTGGCAATGTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCCAGTATGAC	479
QY	141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160
Db	480	AGTGAGAGATCCCTAGGGGATACGCTCCCTTCAGTCAACCAACAGCCAGCATGTCAAGAGAA	539
QY	161	IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg	180
Db	540	ATCCCTGGAGCTTCGCTGATCATCATGATGTCCCTCAGGGGAAATCATCAGCAGACGT	599

QY	181	AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle	200
Db	600	GCTCCGCTTCCCTATGTGAATCATTCACCAATCCGTCAGGGAGTGTCTCCGGCAGTATT	659
QY	201	GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyVala	220
Db	660	GGGAATGTTCCTGGAAAGAGAGTGTGCTGGAAATGAAGCAGGCAAGGTGCG	719
QY	221	IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle	240
Db	720	ATTCCCATGACTAATGGGACAGCATTGTCCCTCTGAAGTCCGCGAGCTACTGACATC	779
QY	241	AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro	260
Db	780	GATGCATCTACTGATACACATGGAAGACGCTTTTACTGAATGATGAATCCCGACCT	839
QY	261	LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal	280
Db	840	CTATCTAGAAAGTCCCATTTGCTTCTCCAAATAAATCCCTACAGATGGTCATTGTT	899
QY	281	LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn	300
Db	900	CTGCGGTGTGTTCTTAGCATCTTCCTGCACCTACCGTCTCACAAATCTGTGCGTAAT	959
QY	301	AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle	320
Db	960	GCATACCCACTGTGGCTTTTATCTGTATATGTGAGATTTGGTTGCTTATCTCTGATA	1019
QY	321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
Db	1020	CTGGATCAGTTCGGAAGTGTGTTCANTCAACCGGAGACCTACCTTGATAGACTGGCT	1079
QY	341	LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer	360
Db	1080	TTAAGGTATGACCGAGAGGTGAACCGTCTCAGTGTGCTGCTGTGCATATTGTGCTAGT	1139
QY	361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380
Db	1140	ACAGTCACGCCCTTGAAGGAGCCACCTATCGTCACTGCCAACACTGTGCTATCCATCTT	1199
QY	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1200	GCTGTTGATTATCCCGTGGACAGGTCTCTTGCTATGTATCTGATGACGAGCTTCAATG	1259
QY	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
Db	1260	CTGACTTTTGACGCAATGGCTGAGACTTCAGAGTTTGTAGGAAATGGTACCATTTGTG	1319
QY	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1320	AAGAAGTATGACATTGAACCCAGAGCTCCGAGTTCCTTTTCTCCAGAAATTTGATTAC	1379
QY	441	LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr	460
Db	1380	CTGAAGACAAAGTCCAGCCTTCATTGTTTAAAGCCCGCGGCCCATGAAGAGAAATAT	1439
QY	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1440	GAAGAAATTAATAATCAGATAAATGCCCTAGTTTCTAAGGCATTTGAAATGCCCGAGAA	1499
QY	481	GlyTrpIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly	500
Db	1500	GGATGGATCATGCAAGATGGCACCATGGCCAGGAAACAATACGAGGATCATCTCTCGA	1559
QY	501	MetIleGlnValPheLeuGlyHisSerGlyLysLeuAspThrGluGlyAsnGluLeuPro	520
Db	1560	ATGATTGAGTTTCTTGTGTCACAGTGTGGCTTGATCTGAGGGTAAATGAGCTCCCC	1619
QY	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly	540
Db	1620	CGTTTAGTTTATGTCTCGTGAAGAGCGCTCTGGGTTCCAGCACCAAGAAGAGCTGCT	1679
QY	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGluTyrMetLeuAsn	560

```

Db      1680  |||||GCAATGAAATCCCTTGTCTGTGCTCTCAGCTGTCTTACTAATCGACATACATGTTGAAT 1739
QY      561  LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db      1740  CTTGATTGTGATCACTACATCAACAACAGCAAGGCTGTCCGAGAAGCTATGTGCTTCCTA 1799
QY      581  MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db      1800  ATGGATCCAAACCTAGTCCGCAAGTCTGHTATGTGCAGTTCCTCCACAAAGGTTTGATGGG 1859
QY      601  IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db      1860  ATTGATAGGAATGATCGATATGCAAAACAGAACACACTGCTCTTTTGTATTAATCTTGAGG 1919
QY      621  GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db      1920  GGCCTTTCAGCGCAATTCAGACCAAGTTTATGTGGGAAGTGGTGTGTCTTTCACAGAAACA 1979
QY      641  AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db      1980  GCTATCTATGGTTATGAGCCCCCAATTAAAGCGAAGACGAGTTCTTTGGCATCACA 2039
QY      661  CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db      2040  TGTGGGGCAAGAAGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAGAGTCG 2099
QY      681  AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db      2100  AACAAAGCATGTGGACAGTTCTGTTCAGTATTCAATCTCGAAGACATAGAGAGGGTGT 2159
QY      701  GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db      2160  GAAAGTGTCTGGTTTGTATGATGAGAAATCAGTTCTCAUGTCTCAAATGAGCTTAGAAG 2219
QY      721  ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
Db      2220  AGATTGTGGCCAGTCAGCAGCAATTTGTGCTCCACTCTGATGAAATATGTTGTTCT 2279
QY      741  GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db      2280  CAGTCTCCACTCCAGAAATCTCTTTGAAGAAGACTATCCATGTCAATAAGTTGTGCTAT 2339
QY      761  GluAspLysSerGluTrpGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db      2340  GAGGCAAGTCTGAATGGGAACTGAGATTGTTGGATCTATGATCTGTCCACAAAGAT 2399
QY      781  IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db      2400  ATTCTTACTGGATTCAAGATGCACGAAGGCTGGCGTTTCAGTCTATTGCAATGCCAAG 2459
QY      801  ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db      2460  CGCCAGCTTTCAAGGGATCTGCCCCCAATCAATCTTTTCAGATCTGTCCAGAAAGTGTG 2519
QY      821  ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
Db      2520  CGGTGGGCTCTCGTTCGTGTAATCTTTTCAGCCGGCATTTGCCCTTATGGTATGGC 2579
QY      841  TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db      2580  TACGAGAGCGCGCTCAAGTCTCTCGAGAGATTGGCTTACATCAACACCACTATTACCCA 2639
QY      861  LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db      2640  CTAACCTCTCTCCGCTCTCTAGTCTATTGTATATTGCGCTGTATCTGTCTGCTCACTGCA 2699
QY      881  LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db      2700  AAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGATCTGTGTTCAITGGCTCTTCTCT 2759
QY      901  SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920

```

```

Db      2760  TCAATTTTCCCACTGGTATCCTTGAGATGAGGTGGAGTGGTGTGGCATTTGACGAGTGG 2819
QY      921  TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db      2820  TGCAGGAATGAACAGTCTTGGGTCAATGGAGGTATCTCTGCACATCTGTTTGGCGTCTT 2879
QY      941  GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db      2880  CAGGGTCTTCTGAAGGTGCTTGCCTGATCGCGGTATCGACACCAACTTCACTGTCAACG 2939
QY      961  AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db      2940  AATGACGAAGAGCGCACCTTGTCTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATC 2999
QY      981  ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db      3000  CCTCCGACGACCACTTTTGATCAATTAACATGGTGTGGTGTGGTCTGGCACCCTCTACGCC 3059
QY      1001  IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db      3060  ATCAACAGTGGTTTACCAATCATGGGGCGCGCTCTTTGGGAAGCTCTCTTTGGCTTCTGG 3119
QY      1021  ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db      3120  GTGATTGTTCACTTATACCAATTCCTCAAGGGTCTTATGGGAGGCAACCGCACCCG 3179
QY      1041  ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db      3180  ACGATTGTCATCGTCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3239
QY      1061  ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db      3240  GTTGATTCATTCACTACCGCTCTCGCTGGCCCAAAATATCCAAACCTGTGGCATCAACTGC 3299

```

RESULT 2

AK072356

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:J023059102, full

insert sequence.

ACCESSION

AK072356

VERSION

AK072356.1 GI:32982379

KEYWORDS

FLI CDNA; CAP trapper.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartioideae; Oryzaceae; Oryza.

REFERENCE

1

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-Length cDNA Project Team::

Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,

Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,

Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,

Science Genome Sequencing & Analysis Group, Otsu,Y., Murakami,K.,

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,

Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,

Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,

Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,

Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN::

Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,

Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,

Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Oka,Y.,

Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,

Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

MEDLINE

22752273

PUBMED

12869764

REFERENCE

2

(bases 1 to 3954)

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,

Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,

Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Ikeda, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nakikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, F., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K. Genomic Sciences Center and Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers

```

1..3954
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023059102"

```

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3954
Score:	5452.50	Matches:	999
Percent Similarity:	97.22%	Conservative:	52
Best Local Similarity:	92.41%	Mismatches:	29
Query Match:	94.37%	Indels:	1
DB:	8	Gaps:	1

US-09-900-237A-30 (1-1080) x AK072356 (1-3954)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20

Db	437	ATGACGGCGACGCGGATGCCGTGAAGTCGGGAGGCACGGGAGCGGCGAGGCGTCCAG	496
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyVasPValPheThrAlaCysAspVal	40
Db	497	ATCTGGCGGACGGCGTGGGACGACGCGGGAGGGGACGCTGTTCCGCCCTCGCAGCTC	556
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlyThrGlnAla	60
Db	557	TGCGGGCTCCCGGTGTGCGCGCCCTGCTACGAGTACGAGCGCAAGGATGCACCCAGCT	616
QY	61	CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
Db	617	TGCCCCCAGTGGCAAGACCAAGTACAAAGCGCCCAAGCGGGAGCCCGCGATCCCTGGGAG	676
QY	81	GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
Db	677	GAAGCGGAGGATGACTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	736
QY	101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly	120
Db	737	GCCGACCCAGAGCAGAAAGATTGCTGATAGATGCGCAGTTGGCGCATGAATGCTGGGGT	796
QY	121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
Db	797	GGTGGAGACGTCGCGCGTCCCAAGTATGACAGTGGCGAGATCGGGCTCACCAGATGAC	856
QY	141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160
Db	857	AGTGGCGAGATCCCTGGGGATACATCCCTTCAGTCACTAATAGCCAGATCTCGGGAGAA	916
QY	161	IleProGlyAlaSerProAspHisHisMetSerProThrGlyAsnIleSerArgArg	180
Db	917	ATCCCTGGAGCTTCCCTGATCATCATGATGATGATGATGATGATGATGATGATGATGATG	976
QY	181	AlaProPheProThrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle	200
Db	977	GCTCCATTCCTTATGTGAACCATTCACCAAAATCCATCAAGGGAGTTCTCTGGCAGCAT	1036
QY	201	GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla	220
Db	1037	GGAAATGTTGCTGGAAAGAAAGAGTGTGCTGGTGGAAATGAAGCAGGACCAAGGGAGCA	1096
QY	221	IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle	240
Db	1097	ATTCCTCATGACCAACGGGACCAAGCATTCGCCCTTCTGAAGGTCGGGAGTTGGTGATATC	1156
QY	241	AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro	260
Db	1157	GATGCATCCACTGATTACAAATATGGAAGATGCCCTTACTGAATGATGAATCACTCCGAGCT	1216
QY	261	LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal	280
Db	1217	CTCTCTAGAAAGTTCCTTCCTTCATCCAGATAAATCCCTACAGAATGGTCATTGTT	1276
QY	281	LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn	300
Db	1277	CTGGGTTTGGTTGTTCTTAAGCATTTTCCTTCATCCAGTATACCAATTCCTGTGCGCAAT	1336
QY	301	AlaTyrProLeuTyrLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle	320
Db	1337	CGGTATCGCTCTGGCTTTTATCTGTTATATGAGATTTGGTTGCTTCTCTCTGGATA	1396
QY	321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
Db	1397	TTGATCAGTTCCCGAGTGGTTCCCAATCAACCGTGAACCTACCTTGATAGGTGSCA	1456
QY	341	LeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheValSer	360
Db	1457	TTGAGGTATGACAGAGAGGTGACCAATCTCAGTTGGCTGCTGTTGATTTTGTTCAGT	1516
QY	361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380

Db	1517	ACAGTCGACCCCATGAAGAACTCTCTTTACTGCCAATACCGTGTCTCCATTCTT	1576
Qy	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1577	GCTGTTGATTACCCAGTGACAGAGGCTCTTGCTATGATCTGACGATGGTCTGCAATG	1636
Qy	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
Db	1637	CTGACTTTTGATGTCATTGCTCAGACTTCAGAGTTTGCTAGAAAGTGGGTTCCATTCGTT	1696
Qy	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1697	AAGAAGATATAACATTGAGCAAGAGCTCTGAGTGGTACTTCTCCAGAAAATCGATTAC	1756
Qy	441	LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr	460
Db	1757	TTCAAAGACAAAGTCCACCCTTCATTGTTAAAGACCGCTGCGCCATGAAGAGAAATAT	1816
Qy	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1817	GAAGAATTCAAAGTTAGGATAAATGGCCTTGTTGCTNAGGCACAGAAATGCCGAGGAG	1876
Qy	481	GlyTrpIleMetGluAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	500
Db	1877	GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATACTAGGACCATCCTGGA	1936
Qy	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1937	ATGATTGAGTTTTCCTTGGTCACAGTGGTGGCTTGATACCGGAGGTAATGAGTTCCTC	1996
Qy	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly	540
Db	1997	CGCCTGGTCTACGTATCTCGTAGAAGCGTCTGGGTTTCAGCACCAAGAGAGCTGGT	2056
Qy	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560
Db	2057	GCCATGAATGCTCTTGTGTCGTCTCAGCTGTGCTTACCAATGGACAGTACATGTTGAAT	2116
Qy	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580
Db	2117	CTTGATTGTGATCACTACATCAACACAGCAGAGGCTCTGCGGGAAGCTATGTGCTCCTT	2176
Qy	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600
Db	2177	ATGGATCCAAACCTAGGAAGAGTGCTGTTATGTTTCAGTTCCCAACAAAGTTTCATGTT	2236
Qy	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620
Db	2237	ATTGATAGGAATGATCGATATGCGAACAGGAACACTGTGTTTTTCGATATTAACCTTGAGG	2296
Qy	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640
Db	2297	GCTCTTGATGCAATCCAAAGGACAGTTTATGTGGGAACCTGGTTGTGTTATCAACAGAACT	2356
Qy	641	AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGly---PheLeuAlaSer	659
Db	2357	GCTCTATATGTTATGATACCCCAATTAAGCAGAGAGAGAGGAGTTCTTGTCTATCA	2416
Qy	660	LeuCysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys	679
Db	2417	CTATGTGGGGCGAGGAAGAGGCAAGCAAGTCAAGAAAGAGAGTTCCGACAAAGAAG	2476
Qy	680	SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGly	699
Db	2477	TCAACACAGCAGTGAGGAGTGCTGTGCCAGTTTTCATCTTGAGATATAGAGAGGGGT	2536
Qy	700	ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu	719
Db	2537	GTTGAAGGTGTGGATTTCGATGATGAGAAATCACTTCTTATGTCTCAATGAGCTTGAG	2596
Qy	720	LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal	739
Db	2597	AAGAGATTTGGCCAGTCTGCAGCGTTTGTTCCTCCACTCTCATGGAATATGGTGGTGT	2656
Qy	740	ProGlnSerSerThrProGlnSerLeuLysGluAlaIleHisValIleSerCysGly	759
Db	2657	CCTCAATCTGCAACCCCAAGATCTCTTTTGAAGAAGCTATCCATGTGATAAGTTGTGGC	2716
Qy	760	TyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu	779
Db	2717	TATGAGGACAAACCGAATGGGAACTGAGATTGGGTGGATCTACGGTTCGGTGACAGAA	2776
Qy	780	AspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro	799
Db	2777	GATATTCTCAGTGGATTCAAGATGTCATGCGAGGCTGGAGATCAATCTACTGCGATGCC	2836
Qy	800	LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal	819
Db	2837	AAGCCGCCAGCTTTCAAGGGGTCTGCTCCTATCAATCTTTTCAGATCGTCTTAACCAAGTG	2896
Qy	820	LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr	839
Db	2897	CTTCCGTGGGCACTTGGTTCTGTTGAAATCTTTTCAGTCGCCATTTGCCCATATGGTAC	2956
Qy	840	GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr	859
Db	2957	GGCTATGGAGGAGCGCTTAAGTTCTTTGGAGAGATTGGCTACATCAACACCACTATTAT	3016
Qy	860	ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr	879
Db	3017	CCATTGACATCGATCCGCTTCTCATATACTGTGTTTGGCTGCTATCTGTTGCTCACT	3076
Qy	880	GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe	899
Db	3077	GGAAAGTTTCATCTCCAGAGATTAGCAACTTGTAGTATTGTTGTTCTCTCTCTTC	3136
Qy	900	LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu	919
Db	3137	ATTTCAAATTTTGGCCAGTATCCTTCAGATGAGTGGAGTGGTGTGGCATCATGAG	3196
Qy	920	TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal	939
Db	3197	TGGTGGAGAAATGAACAGTCTCTGGGTTATTGGAGGTATATCTGCGCATCTTTTTCGGTC	3256
Qy	940	PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys	959
Db	3257	TTCCAGGGTCTCTCAAGTGCTTGTGCTGATTCGACACCAATTTTCACTGTCACTCAAG	3316
Qy	960	AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu	979
Db	3317	GCTTCTGATGAAGATGGCGACTTTTGTGCTGAGCTCTACATGTTTCAAGTGGACACGCTTCTC	3376
Qy	980	IleProProThrThrIleLeuIleIleAsnMetValGlyValAlaGlyThrSerTyr	999
Db	3377	ATCCACCGCAGCAGCATCTTGATCATTAACCTGGTTCGGTGTGTTGTTGTTGTTCTCATC	3436
Qy	1000	AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe	1019
Db	3437	CGGATCAACAGCGCTACCAATCATGGGACCGCTCTTTGGCAAGCTCTTCTTGGCTTC	3496
Qy	1020	TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr	1039
Db	3497	TGGGTGATTGTCCTTGTATCCCTTCTCAAGGGTCTTATGGGTGGGCAACCGCACT	3556
Qy	1040	ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal	1059
Db	3557	CCGACCATCGTGTGTTGGGCAATCTTCTGCTTCGATCTTCTCATTTGCTGTGGGTT	3616
Qy	1060	ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn	1079
Db	3617	CGCATCATTCATCCACCCCGGTGTCACCGGCCAGATATCCCAAAATATGTGGCATCAAC	3676
Qy	1080	Cys 1080	
Db	3677	TGC 3679	

```

RESULT 3
AX653232          3264 bp      DNA      linear      PAT 22-MAR-2003
LOCUS             3102 from Patent WO03000898.
DEFINITION
ACCESSION         AX653232
VERSION           AX653232.1  GI:29156046
KEYWORDS
SOURCE            Oryza sativa
ORGANISM          Oryza sativa
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
1. Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
  Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
  Plant genes involved in defense against pathogens
  Patent: WO 0300898-A 3102 03-JAN-2003;
  Syngenta Participations AG (CH)
FEATURES
    source
    1..3264
    /organism="Oryza sativa"
    /mol_type="unassigned DNA"
    /db_xref="taxon:4530"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3264
Score:          5434.50      Matches:      998
Percent Similarity: 96.60%      Conservative: 52
Best Local Similarity: 91.81%      Mismatches: 30
Query Match:      94.06%      Indels: 7
DB:                6      Gaps: 2

US-09-900-237A-30 (1-1080) x AX653232 (1-3264)
QY      1 MetAspGlyAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB      1 ATGGACGGCGACGGGAGTGGCGTGAAGTGGGAGGACGGGAGCGGCGGCGTGGCCAG 60
QY      21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB      61 ATTCGGCGGACGGCGTGGGACGACGCGGCGGAGCGGCGTGTTCGGCGCTCGGAGTC 120
QY      41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
DB      121 TGGCGGCTTCGGGTGGCGCCCTGCTACGAGTACGAGCGCAAGGATGGCACCCAGGCT 180
QY      61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB      181 TGCCCCCAGTGCAGACCAAGTACAGCGCCACAGGGGAGCCCGCGGATCCCGTGGGGAG 240
QY      81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB      241 GAAGGCGGAGGATACATGATGCTGATGATGTCTAGTACTACAACTACCCCTGCATCTGGCAGT 300
QY      101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
DB      301 GCGGACAGAGAAGCAGAAAGATTGCTGATAGGATGGCGAGTGGCCCAATGAATGCTGGGGGT 360
QY      121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB      361 GGTGGAGACGTGGCGCTCCCAAGTATGACAGTGGCGGAGATCGGGCTCACCAAGTATGAC 420
QY      141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB      421 AGTGGCGAGATCCCTCGGGGATACATCCCTTCAGTCACTAATAGCCAGATCTCGGGGAA 480
QY      161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
DB      481 ATCCCTGGAGCTTCCCTGATCATATATGATGTCTCTACCGGAAACATTTGGCAAGGT 540
QY      181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200

```

```

541 GCTCCATTCCCTATGTGAACCAATTCACCAATCCCAAGGGAGTTCTCTGGCAGCAT 600
QY      201 GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyVala 220
DB      601 GGAATGTGCTGGAAAGAAAGATTGATGCTGGAACTGAACTGAAAGGACCAAGGAGCA 660
QY      221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB      661 ATTCCCATGACCAACGGGCAAGCATTCCTCTGAAGGTGGGGAGTTGGTGATATC 720
QY      241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGlnThrArgGlnPro 260
DB      721 GATGATCCACTGATTACAAATATGAAGATGCTTACTGAATGATGAAACTCGCCAGCCT 780
QY      261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB      781 CTCTCTAGAAAAGTTCCTCTTCTCATCCAGAAATAAATCCCTACAGATGTGTCTGTT 840
QY      281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
DB      841 CTGGGTTCGTGTTCTTAAGCATTTTCTTCACTACCGTATTACGAATCCTGTGCGCAAT 900
QY      301 AlaTyrProLeuTyrLeuSerValIleCysGluIleTyrPheAlaLeuSerTyrIle 320
DB      901 GGTATCCGCTCTGGGCTTTTATCTGTATATGTGAGATTGGTTCCTTTGTCTCGATA 960
QY      321 LeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
DB      961 TTGGATCAGTTCGGGAGTGGTTCCAATCAACCGTGAAACCTACCTTGATAGCTGGCA 1020
QY      341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB      1021 TTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGTCTGCTTGACATTTTGTGAGT 1080
QY      361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB      1081 ACAGTCGACCCCAAGGAACCTCTCTGTGTACTGCCAATACCGTGTGCTCCATCTTT 1140
QY      381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
DB      1141 GCTGTTGATTACCGAGTGGACAGGTCCTCTTGCTATGTATCTCAGCATGCTGCTCAATG 1200
QY      401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheVal 420
DB      1201 CTGACTTTTGATGCATTTGGCTGAGACTTCAGAGTTTGTAGAAAAGTGGGTTCCTTCGTT 1260
QY      421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
DB      1261 AAGAAGTATAAATTCAGCCAGAGCTCTCTGAGTGGTACTTCTCCCAAGAAATCGATTAC 1320
QY      441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
DB      1321 TTGAAGACAAAGTCCACCTTCATTTGTTAAAGCCGCTGTCATGAAGAGAGAAATAT 1380
QY      461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB      1381 GAAGAAATCAAGTAGGATAAATGGCTTTGTTGCTAAGGCACAGAAAGTCCCGAGGAG 1440
QY      481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
DB      1441 GGAATGGATATCAAGATGGCACACCATGGCCAGGAAACAATACTACGGAGGTAATGAGCTCCC 1500
QY      501 MetIleGlnValPheLeuGlyHisSerGlyLysLeuAspThrGluGlyAsnGluLeuPro 520
DB      1501 ATGATTCAGGTTTCTTGTGTGTCAGTGGTGGCTTGTATACCGAGGGTAATGAGCTCCC 1560
QY      521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
DB      1561 CGCCTGTGTACGTATCTCGTGAAGCGCTCTGGGTTTTCAGCACCAACCAAGAGGCTGGT 1620
QY      541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB      1621 GCCATGAATGCTCTTGTTCGTGCTCAGCTGTGCTTACCAATGGACAGTACATGTGAAT 1680

```


QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValAlaArgGluAlaMetCysPheLeu 580
 Db 1681 CTTGATTTGATCACTACATCAACACAGCAAGGCTCTCGGGAAGCTATGCTTCCTT 1740
 QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
 Db 1741 ATGGATCCAAACCTAGGAAGAGTGTCTGTTATGTTCAAGTCCCAAGAGGTTCCATGGT 1800
 QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
 Db 1801 ATTGATAGCAATGATCGATATCGAACAACAGAACTACTGTGTTTTCGATATTAACTTGAGG 1860
 QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
 Db 1861 GGTCTTGATGGCATCAAGGACAGTTTATGTGGGAAGCTGGTGTGTTATCAACAGAAT 1920
 QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGly---PheLeuAlaSer 659
 Db 1921 GCTCTATATGGTTATGAACCCCAATTAAGCAGAGAAGAGGAAGTTCTTGTTCATCA 1980
 QY 660 LeuCysGlyGlyLysLysLysAlaSerLysLysLysArgSerSerAspLysLysLys 679
 Db 1981 CTATGTGGGGCAGGAAGAGCAAGTCAAGAGAGAGAGTTCGGACAAGAGAG 2040
 QY 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
 Db 2041 TCAACAAGACAGCTGGACAGTCTGTCCAGTTTTCATCTTGAAGATATAGAGAGGT 2100
 QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
 Db 2101 GTTGAAGTGTGGATTTCGATGATGAGAAATCACTTCTTATGCTCAATGAGCTTGAG 2160
 QY 720 LysArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
 Db 2161 AAGAGATTGGCCAGCTGCAGCGTTTGTGCTCCACTCTCATGGAATATGGTGGTGT 2220
 QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
 Db 2221 CTTCAATCTGCAACCCCAAGATCTCTTTGAAGAAGCTATCCATGTGATAAGTTGTGCG 2280
 QY 760 TyrGluAspLysSerGluTyrGlyThrGlu-----IleGlyTyrIle 773
 Db 2281 TATGAGGACAAGCCGAATGGGAATCGAGGCTATGCAATGATGGTGGATC 2340
 QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
 Db 2341 TACGGTTCCGTGACAGAAGATATCTCACTGGATTCAAGATGCATCGCGAGGCTGGAGA 2400
 QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
 Db 2401 TCAATCTACTGATGCCCAAGCCCGCCAGCTTTCAAGGGGTCTGCTTCAATCAATCTTCA 2460
 QY 814 AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg 833
 Db 2461 GATCGTCTTAACCAAGTGTCTCGGTGGGCACCTTGTTCTGTGAAATCTTTTCAGTCGC 2520
 QY 834 HisCysProLeuTyrTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
 Db 2521 CATGTCCCATATGTACGGCTATGAGGACGCCCTTAAGTTCTTGGAGAGATTTGCCTAC 2580
 QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuLeuValTyrCysIleLeuPro 873
 Db 2581 ATCAACACCATATTATTCATGTACATGATGATCCCGCTTCTCATATCTGTTGCTT 2640
 QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
 Db 2641 GCTATCTGTTGCTCACTGGGAAGTTTCATATCCACAGAGATTAGCAACTTGTCTGATG 2700
 QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer 913
 Db 2701 TGGTTTCATCTCTCTCATATTTCAATTTTGGCACTGGTATCTTGGATGAGTGGGTGAGT 2760

QY 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
 Db 2761 GGTGTTGGCATCGATGAGTGGTGGAGGAATGAACAGTTCTTGGGTTATTGGAGGTATATCT 2820
 QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
 Db 2821 GCGCATCTTTTTCGGCTCTTCCAGGCTCTCTCAAGGTGCTTGTCTGGTATCGACACCAAT 2880
 QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
 Db 2881 TTCACTGTCACTCAAGGCTTCTGATGAAGATGGCACTTTTGTCTGAGCTTACATGTTTC 2940
 QY 974 LysTyrThrThrLeuLeuLeuProThrThrIleLeuIleLeuAsnMetValGlyVal 993
 Db 2941 AAGTGGACAACGCTTCTCATCCACGACGACCACTTTGATTCATTAACTGTGTCGGTGT 3000
 QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly 1013
 Db 3001 GTTGTCTGTATCTCATACGCGATCAACAGCGCTACCATTCATGGGACCGCTCTTTGGC 3060
 QY 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
 Db 3061 AAGCTCTTCTTTCGCTTCTGGGTGATTGTCCACTTGTACCCCTTCTCAAGGCTCTTATG 3120
 QY 1034 GlyArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIle 1053
 Db 3121 GGTCCGCAAAACCGCACCTCCGACCATCGTTGTGTGGCAATCCTTCTGCTTCGATC 3180
 QY 1054 PheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
 Db 3181 TTCATATTGCTGGTTCGCATCGATCCATCCACCCGTTGTCACCGCCCGCCAGATACC 3240
 QY 1074 GlnThrCysGlyIleAsnCys 1080
 Db 3241 CAAACATGTGGCATCAACTGC 3261

RESULT 4
 AF200533
 LOCUS Zea mays cellulose synthase-9 (Cesa-9) mRNA linear PLN 31-AUG-2000
 DEFINITION Zea mays cellulose synthase-9 (Cesa-9) mRNA, complete cds.
 ACCESSION AF200533
 VERSION AF200533.1 GI:9622889
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 3795)
 AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
 Xoonostle-Cazares,B. and Delmer,D.P.
 TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene
 family
 JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
 MEDLINE 20398328
 PUBMED 10938350
 REFERENCE 2 (bases 1 to 3795)
 AUTHORS Dhugga,K.S. and Helentjaris,T.G.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
 International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
 FEATURES
 Location/Qualifiers
 1..3795
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /note="identified and isolated from the maize genome
 database cDNA collection of Pioneer Hi-Bred International,
 Inc"
 1..3795
 /gene="Cesa-9"
 238..3477
 /gene="Cesa-9"
 gene 1..3795
 CDS 238..3477

/note="CESA-9"
 /codon start=1
 /product="cellulose synthase-9"
 /protein_id="AAF89969.1"
 /db_xref="GI:9622890"

/translation="MEGDADGVKSGRRGGQVCQICDGVGTGAEGDVFTACDVCGFP
 VRCPCYERKDDQACPOCKNKYKRHKGSPIALRGEGDDTDADDSDFNPASGNDD
 QOKKIAADRMKSRMNAAGSGDVGPRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGE
 IPGASPDHMMSPGTGNGIRAPDPYMNHSSNPSREFSGVGNVAKRVDGKMKKODK
 GTIPMNGHSIAPSEGRGVGDDIADSDYNMEDALLNDETRQPLSRKVPKLPSSRINPR
 MWLVLRIVLSIFLHRYINPNIAYPLMWSLICEIFALSWILDLQFPRWFPINR
 YDLRLAIRDYREGEPSQLAAVDIFVTDPMKEPPLVANTVLSILAVDPYDKVSCY
 VSDGAAMLTFDALASTFEARVFKVKNIEIPRAPEYFSQKIDYLKDKVHPSEFV
 KORRAMRYEEFKIRVNGLVAKAQVPEGIMODGTPWGNNTDRHPGMITOVFLGH
 SGGLDTEGNEPLVLVVSREKRPFOHKKAGAMNALVRVSAVLNNGQYMLNDCDHY
 INNSKALRAMCFMPLDNLGRSVYVQFPQREDIDRNDRYANRNVFEDINLRGLDG
 IGPVTVTGCVFNRLIYGPPIKQKGGFSLSCGGRKKGSKSGKSDKRSQKH
 VDSVPVFNLEDEEGVEGAFDEKLSMQMSLEKRFQSAAFVASTIMEYGGVFQ
 SATPESLKEAITHVISCGYEDKIEWTEIGWISVTEIDILTGFKHARGWRSIYCMP
 KPAPFKGSAPINLSRLNOLVRWALGSVEILFRHCLPWYGGRLKFLERFAYINTT
 IYPLTSLPLLIYCILPAICLLTGKFTIPEISNFSIWFISLRFISFATGILEMRWGV
 GIDEMWNEQFWVIGGISALHFAVQGLLKVLAGIDTNETVTSKASDEGDPALRYME
 KWTLLIPTTILLINLGVVAGISAINSGVQSWGLFKGLFFAFWVIVHLVPLFKG
 LMRQNRNTRPTIVVWAILLASIFSLWVRIDPFTNRVTPGDTNRTCGINC"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3795
 Score: 5423.50 Matches: 996
 Percent Similarity: 96.48% Conservative: 46
 Best Local Similarity: 92.22% Mismatches: 37
 Query Match: 93.86% Indels: 1
 DB: 8 Gaps: 1

US-09-900-237A-30 (1-1080) x AF200533 (1-3795)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 Db 238 ATGGAGGGCGACCGGAGCGCGTGAAGTCGGGAGCGCGGGGAGCGGAGGTGTGCCAG 297
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 Db 298 ATCTGGCGCGATGCGCGTGGGCACCTACGCGCGAGGAGAGCGTCTTCACCGCCTCGCAGCTC 357
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
 Db 358 TGGCGGTTCCGGTGTGCGCCCTCTACGATACGAGCGCAAGCGGAGCGGACACAGCG 417
 QY 61 CysLeuGlnCysLysThrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 Db 418 TGCCCCCAGTGTCAAAAACAAAGTACAAAGCGCCAAAGGGAGTCCAGCGATCCGAGGGAG 477
 QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 Db 478 GAAGGAGACGATGATGCGGATGATGCTAGGAGCTTCAACTACCGTGCATCTGCGCAAT 537
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
 Db 538 GAGCACCAAGACAGAGATTGCTGACAGATCGCGAGCTGGCGCATGATGTCGGGGC 597
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 Db 598 AGCGGGGATGTTGGCGCGCCCAAGTATGACAGTGGTGGATCGGGCTTACCAAGTACGAC 657
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 Db 658 AGTGGTGAGATCCCTCGGGATACATCCCGTCAGTCAGTACAGCGGAGATTCGGGAGAA 717
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 Db 718 ATCCCTGGTGTCTCCCTGACCATCATATGATGTCTCTACTCGGGAACATTGCGAGCGC 777
 QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200

Db 778 GCCCATTTCCCTATATGAATCATTCATAATCCGTCAGGGAATTCCTCGGTAGCGTT 837
 QY 201 GlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
 Db 838 GGGAAATGTTGCCCTGGAAAGAGAGGGTGTGATGGCTGAAATGAAGCAGGACCAAGGAACA 897
 QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
 Db 898 ATTCCCATGACGNAATGGCACAGCATTTGCTCCTCTGAGGGCCGGGGTGTGGTGATATT 957
 QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuLeuAsnAspGluThrArgGlnPro 260
 Db 958 GATGCATCAACTGATTACCAACATGGAAGATGCTTATTAAACCATGAAATCTGCCAGCCT 1017
 QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 Db 1018 CTATCTAGGAAAGTTCCACTTCTCTCCAGGATAATCCATACAGGATGGTCATTGTG 1077
 QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 Db 1078 CTACGATTGATTGTTCTAAGCATCTTCTTGCACTACCGGATCACAAATCCTGTGCGTAAT 1137
 QY 301 AlaTyrProLeuTyrLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTyrIle 320
 Db 1138 GCATACCCACTGTGGCTTCTATCTGTATATGTAGATCTGTGCTGCTTCTCTCTCGATA 1197
 QY 321 LeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
 Db 1198 TTGGATCAGTTTCCAAAGTGTGTTCCAAATCAACCGGAGACTTACCTTGATGACTGCCA 1257
 QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
 Db 1258 TTAAGGTATGACCGGGAAGGTGAGCCATCTCTAGTTGGCTGCTGTGACATTTTGTCA 1317
 QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
 Db 1318 ACTGTGCGACCAATGAAGGAGCCTCTCTTGTCTACCTGCCAATACCGTGTCTATCCATTCTC 1377
 QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 Db 1378 GCTGTGGACTATCCTGTGGATAAGGTCTCTTGCTATGTATCTGATGATGAGCTGCTATG 1437
 QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheVal 420
 Db 1438 CTGACATTTGATGACTAGCTAGACTTCAGAGTTTGTCTAGAAAATGGGTGCCATTGTT 1497
 QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 Db 1498 AAGAAGTACAACTTGAACCTAGAGCTCTGATGGTACTTCTCCAGAAAAATTGATTAC 1557
 QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
 Db 1558 TTGAAGACAAAGTGCACCTTCATTTGTTAAAGCCCGCGGCCATGAAGAGAGATAT 1617
 QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 Db 1618 GAAGAATTCAAAATTAGGGTAAATGCTTGTGTCTAAGGCACAAAAAGTCCCTGAGGAA 1677
 QY 481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
 Db 1678 GGATGATCATGCAAGATGGCACCATGCGCCAGGAAAAAATAATCCAGGGACCATCTCTGGA 1737
 QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
 Db 1738 ATGATTCAGTTTCTTGGTTCAGTGTGTGCTGTGATGATGAGGGTAAATGAGCTACCC 1797
 QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
 Db 1798 CGTTTGGTCTATGTTTCTCGTGAAAAACGTCTCTGGATTCCAGCATCACAAAGAGCTGGT 1857
 QY 541 AlaMetAsnAlaLeuValValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560

```
Db 1858 GCCATGAATGCTCTTGTCCGCGCTCTCAGCTGTGCTTACCAATGAGCAATACATGTTGAAT 1917
QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db 1918 CTGTGATTGATCACTTACATCAACACAGTAAGGCTCTCAGGAAAGCTATGCTTCCTT 1977
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db 1978 ATGGATCTAAGCTAGGAGAGTGTCTGTATGTTCAAGTTCCTCCAGAGGTTCCATGGT 2037
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db 2038 ATTGATAGAAATGATCATGATATCCCAACAGAAACACCGTGTCTTTTCGATATTAACTTGAGA 2097
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 2098 GGTCTGTATGGCATCCAGGACAGCTTATGTATGGGCACCTGGCTGTGTTTCAACAGAAACA 2157
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2158 GCTCTATATGTTATGACCCCAATTAAAGCAAAAGAGGTTGGTCTTCTGTCACTACTA 2217
QY 661 CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db 2218 TGTGTGGCAGGAAGGAGCAAAATCAAAAGAAAG---GGCTCAGACAAAGAAAAAGTCA 2274
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2275 CAGAAAGCATGTGGACAGTTCTGTGGCAGTATTCAATCTTGAAGATATAGAGAGGAGTT 2334
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2335 GAAGCGCTGGATTGTATGATGATGAGAAATCACTTCTTATGTCTCAATGANGCTTGAGAG 2394
QY 721 ArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
Db 2395 AGATTGGCCAAATCTGCAGCTTTTGTGGTCCACTCTGATGAAATATGGTGGTGTCT 2454
QY 741 GlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2455 CAGTCTGGCAGTCCAGAAATCTCTCTGAAAGAGCTATCCATGTATCAATGTTGTGGCTAC 2514
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db 2515 GAGGACAGATGATGGGAACCTGAGATGGGTGGATCTATGTTCTGTGACGAAAGAT 2574
QY 781 IleLeuThrGlyPheLysMethHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2575 ATTCTCACTGGGTTCAAGATGCACGACGAGGCTGGCGTCTGATCTACTGCAATGCCTAAG 2634
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2635 CGCGCGGCTCTCAAGGAGATCGGCTCCCAATCAATCTCTCAGACCGTCTGAACAGGTGCTC 2694
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2695 CGGTGGGCTCTCGGTTCAATGGAATCTCTTTCAGCGGCAATGCCCCATATGTTACGGG 2754
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2755 TACGGAGGACGCGCTCAAGTCTCTGGAGAGATTCGCTTACATCAACACACCATCTACCCG 2814
QY 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2815 CTCAGTCCCTCCCGCTCTCAATTTACTGTATCTCTGCTGCAATCTGCTGCTCAGGGG 2874
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2875 AAGTTCAATCATCCAGAGATCAGCAACTTCGCTAGTATCTGTTTCACTCTCTCTCAIC 2934
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 2935 TCGATCTTCGCCACGGGTATCTCGAGATGAGGTGGAGCGGCTGGGCATCGACAGTGG 2994
```

```
QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2995 TGGAGAAACGAGCAGTTCCTGGGTCAATCGAGGCACTCCGCCACCTCTTCGCCCTTTC 3054
QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3055 CAGGCGCTCTCAAGGTGCTTGGCGGCATCGACCAACTTCACCGTCACCTCCAGGGC 3114
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle 980
Db 3115 TCGGATGAAGACGGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGACACTTCTGATC 3174
QY 981 ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3175 CGCCCAACACCATCTGATCATCACTCGGTGGCGGTGTTGCGCGCATCTCCCTACGGC 3234
QY 1001 IleAsnSerGlyTyrGlnSerTyrProLeuPheGlyLysLeuPhePheAlaPheTyr 1020
Db 3235 ATCAACAGGGGTACCAAGTCTGGGGTCCGCTCTTCGGCAAGCTCTTCTCGCCTCTGG 3294
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3295 GTGATCGTTCCTGATCCCGTTCCTCAAGGCTCTCATGGGTCTCATGAGTGGCAACCGCAC 3354
QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Db 3355 ACCATCGGTGTCTGGCGCATCTGCTGGCGTCGATCTCTCTCTGCTGGGTTCGC 3414
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3415 ATCGATCGTTCACCAACCGGCTCACTGGCCCGGATCTCGAACCTGTGGCATCAACTGC 3474

RESULT 5
AX338680
LOCUS
DEFINITION
Sequence 5 from Patent WO0179516.
AX338680
ACCESSION
AX338680.1 GI:18129042
VERSION
KEYWORDS
SOURCE
Zea mays
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
Dhugga, K.S. and Helentjaris, T.G.
Maize cellulose synthases and uses thereof
Patent: WO 0179516-A 5 25-OCT-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
1..3799
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
238..3477
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD20545.1"
/db_xref="GI:18129043"
/db_xref="REMBL:CAD20545"
/translation="MEGDAGVKSRRGGGVQICGQGVGTTAEGDVFTACDVCGFP
VCRPYERKDGKQACQCKNKYKRHKSGPAIRGEEDGDDADADADFNTPASGND
QKQIADRMRSWRNAGSGDVGRPKYDSIGLTKYDSGEIPIRGYIPSVNTSISGE
IPGASPDHMMSPGTGNIRRAPFPYMHSSNPREFSGVGNVAKERVDGSKWKQDK
GTIPWNTGTSIAPSEGRGVGDI DASTYNNMEDALLNDETROP LSKVKYPLPSRINPYR
MVIVLRILVLSIFLHYRTNVRNAYPLWLISVCEIWFALSWILDVLPKWPINRET
YLDRLALYRDEGEPSQLAADI FVSFVDPKKEPLTANTVLSILADYFDVKVSCY
VSDGAAMLTFDALAETSEFARKWPFVKVNIIEPAPWFYSOKI DYLKDVHPSFV
KDRAMKREYEEFIRVNLGVAQKQPEEGIMQDGT PWFNNTRHPGMIOVFLGH
SGGLDTGNELPRLVYRSKRKQFQHKKAGAMNALVRVSATVNTGNGMLNLCDDHY
INNSKALREAMCFI LMDPNLGRSCYVOPFORFDGIDRNDR YANRNTVFDINLRGLDG
IQGPVYVGTGCVFNRTALYGYEPPIKQKGGFLSLCGGRKKKSGSKSGSKSKQKH
```


PLTSLPIILYICILPAICLLKFKFIIPISNFASIFWISLFSIFATGILEMRWSGVGI
DEWNRQFVIGISIAHFAVQGLKVLKAGIDNTVTSKASDEGDGDAELVYMKW
TTLLIPPTTILINLVGVAGISVAINSGVQSMPLRGKLFPAFWLVHLYPFLKGLM
GRQNRPTIIVWAILLASIFSLWLRIDFTRVTPDITQTCGINC"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3745
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservations: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 8 Gaps: 2

US-09-900-237A-30 (1-1080) x AF200528 (1-3745)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB 320 ATGGAGGGCGACGCGGACGGCGTGAAGTCGGGGAGCGGTGGCGACAGGTGTGCCAG 379
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB 380 ATTCGGCGGACGCGGTGGGACACACCGCGGAGGGGACGCTTCGCGCGCTGGCAGTC 439
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
DB 440 TCGCGGTTCCGGTGTGCGCCCTGCTACGAGTACGAGCGCAAGGACGCGCAGCGCG 499
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB 500 TGCCCCCAGTGCAGAACCAAGTACAAAGCGCCAAAGGGGAGCGCGCGATCCGTTGGGGAG 559
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB 560 GAAGGAGACGACACTGATGCGGAT-----AGCGACTTCATTACCTTGCACTGGCAAT 613
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
DB 614 GAGGACCAAGCAGAGAAGATTGCGCAGAGAATGCGCAGATGCGCGCATGAACGTTGGGGGC 673
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB 674 AGCGGGATGTTGGTGCAGAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGAC 733
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB 734 AGTGGCGAGATTCTCGGGATACATCCCATCAGTCACTAACACGACAGATCTCAGGAGAA 793
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
DB 794 ATCCCTGGTGTCTTCCCTGACCATCATATGATGTCCTCCCACTGGGAACATTGGCAAGCGT 853
QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
DB 854 GCTCCATTCTCCATGTGAACCAATTGCGCAAAATCCGTCAAGGGAGTTCTCTGGTAGCATT 913
QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220
DB 914 GGGAAATGTTGCTTGGAAAGAGAGGGTTGAGGTGGTGAATGAAGCAGACAGGGGACG 973
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB 974 ATTCCTCATGCAATGGCAGCAACGATTTGCTCCCTCTGAGGGGTGGTGGTGATATT 1033
QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
DB 1034 GATGATCACTGATTACACATGGAGATGCTTATTGAACGACCAAACTCAGACGCT 1093
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB 1094 CTATCTAGGAAAGTCCACTTCTCTCCAGGATAAATCCATACAGGATGGTCAATTGG 1153
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300

DB 1154 CTGCGATTGATTGTTCTAAGCATCTTCTGCACCTACCGTATCACAATCTGTGGCGAAT 1213
QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
DB 1214 GCATACCATTATGGCTTCTATCTGTATATGTGAGATCTGGTTGCTCTTTTCGTGGATA 1273
QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
DB 1274 TTGATCAGTTCCTTAAGTGGTTTCCAAATCAACCGGAGAGCTACCTTGTATAGGCTGGCA 1333
QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB 1334 TTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTTGCTCAGT 1393
QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB 1394 ACAGTCGACCAATGAAGGAGCTCTCTTGTGTCACTGCCAATACCGTGTCTATCCATTCTT 1453
QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
DB 1454 GCTGTGATTACCTGTGGATAAGGTCTCTTGTGTATGTATCTGATGATGGAGTGGCAGT 1513
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
DB 1514 CTGACATTTGATGCACCTAGCTAGAGACTTCAGAGTTTGTAGAAAATGGGTACCATTTGTT 1573
QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
DB 1574 AAGAAGTACAACTTGAACCTAGAGCTCTCTGAATGGTACTTCTCCAGAAAATGATTATAC 1633
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
DB 1634 TTGAGGACAAAGTGCACCTTCATTGTTAAAGACCGCGCGCCATGAAGAGAAATAT 1693
QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB 1694 GAAGAATTCAAAGTAGGTAATGGCTTGTGTTAAGGCACAGAAAATTCCTGAGGAA 1753
QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
DB 1754 GGATGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCMGGGACCATCTCGGA 1813
QY 501 MetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeuPro 520
DB 1814 ATGATTCAGGTTTCTTGGTCACTGGTGGCTTGATCTGAGGCAATGAGCTACCC 1873
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
DB 1874 CGTTTGGTCTATGTTCTCGTGAAGCGTCTCGGATTCAGCATCACAGAAAGCTGGT 1933
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB 1934 GCCATGAATGCTCTTGTTCGTGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAT 1993
QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
DB 1994 CTTGATTGTGATCACTACATTAAACAGTAGGCTCTCAGGGAAGCTATGTCTCTCTT 2053
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
DB 2054 ATGGACCTTAACCTAGGAAGGAGTGTCTGCTACGTCCAGTTTCCCGAGATTTCGATGGC 2113
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
DB 2114 ATTGACGGAATGATCGATATGCCAAGAACACCGTGTGTTTTCGATATTAATCTTGAGA 2173
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
DB 2174 GGTCTTGATGGCATCCAGGACCCAGTTTATGTCCGAACCTGGCTGTGTTTCAACCGAACA 2233
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660

Db 2234 GCTCTATATGTTATGAGCCCCCAATTAAAGCAGAGAGGGTGGTTTCTTGTTCATCACTA 2233
 Qy 661 CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
 Db 2294 TGTGGCGTAGGAAGAGCAAGCAATCAAGAAG--GGCTCGGACAAGAAGATCG 2350
 Qy 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspLleGluGluGlyVal 700
 Db 2351 CAGAAGCATGTGGACAGTTCTGTGCCAGTATTCACCTTGAAGATATAGAGAGGAGTT 2410
 Qy 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
 Db 2411 GAAGCGCTGATTTGACGACGAGAAATCACTTCTATATGCTCAATGAGCGCTGGAGAAG 2470
 Qy 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
 Db 2471 AGATTGGCCAGTCCGAGCGTTTGTTCCTCCACTCTGTATGGAGTATGGTGGTTCCT 2530
 Qy 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaLleHisValLleSerCysGlyTyr 760
 Db 2531 CAGTCCGCACTCCGGAGTCTCTTCTGAAGAAGCTATCCATGTTATAGCTGTGGCTAT 2590
 Qy 761 GluAspLysSerGluTyrGlyThrGluLleGlyTyrLleTyrGlySerValThrGluAsp 780
 Db 2591 CAGGACAGACTGATGGGAACTGAGATCGGGTGGATCTACGGTTCGTGACAGAGAC 2650
 Qy 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
 Db 2651 ATTCTACCGGATTCAGATGACGCGGAGGCTGGCGGTGCATCTACTGTCATGCCCAAG 2710
 Qy 801 ArgProAlaPheLysGlySerAlaProLleAsnLeuSerAspArgLeuAsnGlnValLeu 820
 Db 2711 CGGCCAGCTTTCAAGGGGTCTGCCCCCAATCAATCTTCGACCCGTCGTGACAGGTGCTC 2770
 Qy 821 ArgTrpAlaLeuGlySerValGluLleLeuPheSerArgHisCysProLeuTyrTyrGly 840
 Db 2771 CGGTGGGTCTTGGTTCGTGGAGATCTCTTTACGCGGCACCTGCCCCCTGGTACGCG 2830
 Qy 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
 Db 2831 TAGCGAGGCGGCTCAAGTTCCTGGAGAGATTCGGGTATCATCAACACCACTATCCCG 2890
 Qy 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
 Db 2891 CTCACGTCCATCCCGTCTCTCATCTACTGCATCTCCGCGCCCATCTGCTGTCCACCGA 2950
 Qy 881 LysPheIleMetProGluLysSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
 Db 2951 AAGTTTCATCTCCAGAGATCAGCACTTCGCGAGCATCTGGTTTCATCTCCCTCTTCATC 3010
 Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
 Db 3011 TCGATCTTCGCCACCGGCATCTCGAGATGAGTGGAGCGGGGTGGGCATCGACGAGTGG 3070
 Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
 Db 3071 TGGAGAACGACGAGTCTGGGTGATCGGGGCATCTCCGCGCATCTTCGCGGTCTTCGCGTTC 3130
 Qy 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
 Db 3131 CAGGGCTCTCAGGTGTGGCGCGCATCGACACCACTTCACCGTCACTCCCAAGGC 3190
 Qy 961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuLle 980
 Db 3191 TCGGACGAGGACGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGACGCTCCTGATC 3250
 Qy 981 ProProThrThrIleLeuLleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
 Db 3251 CCGCCCAACCACTCTGTATCAACCTGGTGGCGGTGCTGCCGCGCATCTCCTACGCC 3310
 Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
 Db 3311 ATCAACAGCGGATACCACTGCTGGGGCCCGCTCTTTCGGCAGCTCTTCTTCGCTTCTGG 3370

Qy 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
 Db 3371 GTCATCTGCACCTGTACCCGTTCTCAAGGGCTCTATGGCAGCAGACCGCACCCCG 3430
 Qy 1041 ThrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
 Db 3431 ACCATCTGTCGTCGTGGCCATCTGTGGCGTCCATCTTCTCTGCTGTGGGTTCGC 3490
 Qy 1061 ValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyLeuAsnCys 1080
 Db 3491 ATCAGCCCTTCACACCCGCTCACTGGCCCGGATACCCAGACGTGTGGCATCACTGC 3550
 RESULT 7
 AX652952 3222 bp DNA linear PAT 22-MAR-2003
 LOCUS Sequence 2822 from Patent WO03000898.
 DEFINITION AX652952
 ACCESSION AX652952
 VERSION AX652952.1 GI:29155766
 KEYWORDS
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO 0300898-A 2822 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES
 Location/Qualifiers
 source 1..3222
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3222
 Score: 5169.50 Matches: 950
 Percent Similarity: 94.52% Conservative: 67
 Best Local Similarity: 88.23% Mismatches: 54
 Query Match: 89.47% Indels: 5
 DB: 6 Gaps: 4
 US-09-900-237A-30 (1-1080) x AX652952 (1-3222)
 Qy 6 AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGly 25
 Db 4 GACGTGGCAATCGGGAAGCAA-----TGCCATGCTGCGCAGATCTGCGCGCAGCGC 57
 Qy 26 LeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProVal 45
 Db 58 GTGGGCACGCGCGCGACGCGAGCTCTTCACCGCTCGCAGCTGCGGGTTCGCGGTG 117
 Qy 46 CysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLys 65
 Db 118 TCGCGGCCATGCTACAGTACGAGTACGAGCAGCAGGACGCGCGCGTCCGCGTCAAG 177
 Qy 66 ThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGluGlyAspAspThr 85
 Db 178 ACCAGTACAAGCGCACAAAGGGGCGCGCGATCTTGGGATGAAAGCATGATGTT 237
 Qy 86 AspAlaAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspGlnLysGln 105
 Db 238 GATCGCGATGATGCTAGTGTGAACTATCAACATCCGCGCAACCGACCATTAAGCAC 297
 Qy 106 LysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySerGlyAsnValGly 125
 Db 298 AAGATTGCCGAGAGGATGCTCACCTGGCGCATGAATCTGGGGAGGAATGATGATATGTC 357
 Qy 126 HisProLysTyrAspSerGlyGluLleGlyLeuSerLysTyrAspSerGlyGluLlePro 145

Db 358 CATTTAAAGTATGATAGTGGTGCAGATTGGTCAATCCCAAGTATGACAGTGGTGAATCCCT 417
Qy ArgGlyTyrValProSerValThrAsnSerClnMetSerGlyGluIleProGlyAlaSer 165
Db 418 CGCATAATATATCCCATCGCTCACTCAAGCAGATCTCAGGTGAATTCCTGGAGCGTCC 477
Qy 166 ProAspHisHisMetSerProThrGlyAsnIleSerArgArgAla---ProPhePro 184
Db 478 CCTGAT---CATATGATGTCTCCCGTTGGGNACATTGGCAGACGTGGGCATCCATTTCCC 534
Qy 185 TyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValAla 204
Db 535 TATGTGAACCATTCACCAACCCATCAAGGAGTTCCTGGTAGCCTTGGCAATGTTGCA 594
Qy 205 TrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAlaIleProMetThr 224
Db 595 TGGNAGAGAGAGTAGAGTGGCTGGAAATGAAG---GATAAAGGTGCAATTCCTCCATGGCT 651
Qy 225 AsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThr 244
Db 652 AATGGTACTAGCATTCCTCTCAGAAGTCTGGGGTGGTGACATTTGATGCACTACT 711
Qy 245 GluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLys 264
Db 712 GACTATAACATGGAGGAGCGCTTACTCAATCATGAACACGCCAACCTCTATCGAGAAAA 771
Qy 265 ValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValIleArgLeuVal 284
Db 772 GTGCCAATTTTCGTCATCTCAGATAAATCCGTACAGAGTGTATTTGTGCTCCGATTTGATT 831
Qy 285 ValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeu 304
Db 832 GTCCTATGTATATTTCTTGCACTACCGTATCACAAATCCTGTGGTGAATGCAATCCGCTA 891
Qy 305 TrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPhe 324
Db 892 TGGTTCCTCTCTCATATGTGAGATTGGTTGTCTATCTCTGATGCTTCTGAGTATCAGTTC 951
Qy 325 ProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAsp 344
Db 952 CCCAAGTGTGCCCATCAACCGTGAACCTTACTCGATAGGCTGGCTTTAAGGTATGAC 1011
Qy 345 ArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspPro 364
Db 1012 CGGGAAGGTGAACCATCTCAATTGGCTCCTGTTGACATTTTGTGATGCTGATGCTGATCCT 1071
Qy 365 LeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyr 384
Db 1072 ATGAAGAAGCCCTCCTGTGCTACTGCCAATCTGTGCTCTCCATCCTGCTGTGGATTAC 1131
Qy 385 ProValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAsp 404
Db 1132 CCTGTTGATAAGGTATCTTGCTATGTCTGATGATGGAGCTGCAATGTTAATTTGAT 1191
Qy 405 AlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAsp 424
Db 1192 GCGCTGTGCTGAGACTTCAGAAATTTGTAGAAAATGGGTACCGTTCCTGTAAGGTACAGC 1251
Qy 425 IleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLys 444
Db 1252 ATAGAACCCAGAGCTCCAGAGTGGTACTTTGCACAGAAAATTTGATTACTTTGAAGACAAA 1311
Qy 445 ValGlnProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLys 464
Db 1312 GTTCAGGCTCTTTTGTAAAGATCGTCTGTCGCAATGAGAGGGAATATGAAGAAATTAA 1371
Qy 465 IleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMet 484
Db 1372 GTTCGTGTAAAGTCTCTGCTCCAGGACAGAAAATTTCTCGAGAGGATGGATTATG 1431
Qy 485 GlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal 504

Db 1432 CAAGATGGCACACCTTGGCTGGTAAACAATACCAGAGACCATCTCTGGAATGATTCAGGTT 1491
Qy 505 PheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyr 524
Db 1492 TTCCTTGGTCATAGTGGAGGCTTGATACCGAAGCAATAGCTTCCTCGTTTGTAGTCTAT 1551
Qy 525 ValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAla 544
Db 1552 GTGTCTCGTGAGAAACGTCTCGAATTCACACCATAAAGGCTGTGTATGAATGCA 1611
Qy 545 LeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp 564
Db 1612 CTTGTTCTGTATCAGCTGTCTTACTAATGGACAATACTTGTGTGAATCTTGACTGTGAT 1671
Qy 565 HisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsn 584
Db 1672 CACTACATCAACAATAGCAAGCTCTCCGAGAGGCTATGTGTCTCTTATGTGATCCAAAC 1731
Qy 585 LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn 604
Db 1732 CTAGGAAGGCGTGTCTGTATGTCCCAATCCCTCAGAGGTTTGACGGTATCGATAGAAAT 1791
Qy 605 AspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGly 624
Db 1792 GATCGATATGCAAAACAGGAACACCGTGTCTTTCGATATTAATTTGAGAGGCTCTCGATGGT 1851
Qy 625 IleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGly 644
Db 1852 CTCCAAGGACCACTTTATGTGGAAACGGTGTGTGTCTCAACAGAAACAGCTCTTTATGGT 1911
Qy 645 TyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeuGlyGlyLys 664
Db 1912 TATGAACCCCCCATTAAGCAGAAAGAGCGGAGTACTCTCTTCGCTTTGTGGGGACGA 1971
Qy 665 LysLysAlaSerLysSerLysArgSerSerAspLysLysLysSerAsnLysHisVal 684
Db 1972 AAGAAGACAAAAAAGTCGAAGGAGAGAGACCGAAGAAAGAAAGTCACACAAACATGTG 2031
Qy 685 AspSerSerValProValPheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGly 704
Db 2032 GACAGTCTCTGTGCCAGTTTAAATCTTGAAGATATAGAGGAGGATTTGAAGTCTCTGGA 2091
Qy 705 PheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGln 724
Db 2092 TTTGATGATGAGAAATCACTACTGATGTCTCAATAGAGCTTAGAGAAAAGATTGGTCAA 2151
Qy 725 SerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValProGlnSerSerThr 744
Db 2152 TCTAGTGTTTTGTAGCCCTCCACTCTGATGGAATATGGTGGTGTCTCTCAATCCGCACT 2211
Qy 745 ProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSer 764
Db 2212 CCAGAATCTCTTCTGAAGAGGCCATACATGTATATCAGCTGTGGCTATGAAGACAAAGT 2271
Qy 765 GluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGly 784
Db 2272 GACTGGGAACCTGAGATTGGATGGATCTATGGTCTGTGTACAGAGATATTTCTACTGGA 2331
Qy 785 PheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPhe 804
Db 2332 TTCAAGATGTCATGCACGTGGCTGGCGGTCAATCTACTGCATCCCTAAGCGACGACATTC 2391
Qy 805 LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeu 824
Db 2392 AAGGGGTCTGCCCAATTAATCTTTCGGATCGCTGAAACAGGTGCTTCGGTGGGCTCTT 2451
Qy 825 GlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArg 844
Db 2452 GGTCTGTGGAAATCTCTTTTAGCAGGCATTCCTCCATATGATGTATGGGTACGAGGACGG 2511
Qy 845 LeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeu 864
Db 2512 CTTAAGTCTCTGGAGAGATTTGCTTACATCAACACCACTATTATTCATCCATCGATT 2571

Alignment Scores:

Pred. No.: 0 Length: 4282
 Score: 5165.50 Matches: 949
 Percent Similarity: 94.52% Conservative: 68
 Best Local Similarity: 88.20% Mismatches: 54
 Query Match: 89.40% Indels: 5
 DB: 8 Gaps: 4

US-09-900-237A-30 (1-1080) x AK069196 (1-4282)

Qy	6	AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGly	25
Db	278	GACGGTCGGAATCGGGGAACAA-----TGCCATGTCTGCCAGATCTGGCGCAGCGC	331
Qy	26	LeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProVal	45
Db	332	GTGGCGACGGCGCGGAGCGGAGCTCTTCACGCGCTCGCAGCTCTCGGGTTCGCGGTG	391
Qy	46	CysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLys	65
Db	392	TGCCGGCATGTCTACGAGTACGAGCGCAGCAGCGCAGCGCGCTGCCCGCAGTGCAG	451
Qy	66	ThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGluGlyAspAspThr	85
Db	452	ACCAAGTACAAAGCGCCCAAGGGGAGCGCGCGGATCTTGGGGATGAAAGCGATGATGT	511
Qy	86	AspAlaAspGlySerAspPheAsnTyrProIaSerGlyThrGluAspGlnLysGln	105
Db	512	GATGCGGATGATGCTAGTGTGTAATCCAACTATCCACATCCGCGCAACCGAGGACCAATAGCAC	571
Qy	106	LysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGlySerGlyAsnValGly	125
Db	572	AGANTGCCAGAGGATGCTCACCTGGCGCATGAACTCTGGGAGGAATGATATTTGTC	631
Qy	126	HisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIlePro	145
Db	632	CATTCTAAGTATGATAGTGGTGAATTTGGTCAATCCCAAGTATGACAGTGGTGAATCCCT	691
Qy	146	ArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSer	165
Db	692	CGATATATATCCCATCGCTCACTCACGACGAGATCTCAGGTGAAATTCCTGGAGCGTCC	751
Qy	166	ProAspHisHisMetSerProThrGlyAsnIleSerArgArgAla---ProPhePro	184
Db	752	CCTGAT---CATATGATGTCTCCGTTGGGNACATTCGCAGACGTGGGCATCTCC	808
Qy	185	TyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValAla	204
Db	809	TATGTGAACCATTCACCAAAACCCATCAAGGGAGTTTCTGGTAGCCTTGGCAATGTTGCA	868
Qy	205	TrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThr	224
Db	869	TGGAAGAGAGAGTAGAGTGGCTGGAAATGAAG---GATGAAGTGCATTCCTCCATGGCT	925
Qy	225	AsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThr	244
Db	926	AATGGTACTAGCATTTGCTCTTCAGAGGTCTGGGGTGGTGATGATGATCTACT	985
Qy	245	GluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLys	264
Db	986	GACTATAACATGGAGGAGCGCTTACTGAATGATGAAACACGCCAACCTCTATCAGAAAA	1045
Qy	265	ValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuVal	284
Db	1046	GTGCCAATTTGCTATCCAGAAATTAATCCGTACAGATGGTTATTTGCTCCGATGATTT	1105
Qy	285	ValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeu	304
Db	1106	GTCCTATGATATCTTGCATACCGTATCAAAATCTGTGCGTAATGCATACACCGCTA	1165
Qy	305	TrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPhe	324
Db	1166	TGGTTGCTCTCTGTCATATGTGAGATTTGGTTTGTCTCTATCTCTGGATTCAGTTTC	1225

Qy	325	ProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAsp	344
Db	1226	CCCAAGTGGTCCCAATCAACCGTGAACATTACTCGTGAAGCTGGCTTTAAGGTATGAC	1285
Qy	345	ArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspPro	364
Db	1286	CGGGAAGGTGAACCAATCTCAATGGCTCCTGTTGACATTTTGTCACTACTGTGATCCT	1345
Qy	365	LeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyr	384
Db	1346	ATGAAAGAACCTCTCTCTGTCTCACTGCCAATACTGTCTCTCCATCTCTGTGTGATAC	1405
Qy	385	ProValAspLysValSerCysTyrValSerAspAspGlyValAserMetLeuThrPheAsp	404
Db	1406	CCTGTTGATGAAGGTATCTTGTATGTCTGTATGATGAGCTGCAATGTTAACTTTGAT	1465
Qy	405	AlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAsp	424
Db	1466	GCCTTGTCTGAGACTTCAGAAATTTGCTAGAAAATGGGTACCGTTCTCTGTAAGATACAGC	1525
Qy	425	IleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLys	444
Db	1526	ATAGAACCCAGAGCTCCAGAGTGGTACTTTTGCACAGAAAAATTGATTACTTTGAAAGACAAA	1585
Qy	445	ValGlnProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLys	464
Db	1586	GTTTCAGGCTCTTTTGTAAAGATCTGCTGCATGAGAGGGAATATGAGAAATTTAAA	1645
Qy	465	IleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMet	484
Db	1646	GTTTCGTGTTAATGCTCTCGTCGCCAAGGACAGAAAGTTCTCTGAGGAAGATGATGATTG	1705
Qy	485	GlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal	504
Db	1706	CAAGATGGCACACCTTGGCTGGTAACTAACACAGAGACCATCTCTGGAAGATTTCAGGTT	1765
Qy	505	PheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyr	524
Db	1766	TTCTTGGTCTATAGTGGAGCCCTTGATACCGAAGGCATGAGCTTCTCGTTAGTCTAT	1825
Qy	525	ValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAla	544
Db	1826	GTGTCTCTGAGAAAACGTCTGGATCCAAACACCAACCAAAAGGCTGGTGTATGAATGCA	1885
Qy	545	LeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp	564
Db	1886	CTTGTCTGTATCAGCTGTCTTACTTAATGGCAATACTTGTGTAATCTTGACTGTGAT	1945
Qy	565	HisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsn	584
Db	1946	CATCATCATCAATAGCAAGCTCTCCGAGAGGCTATGTCTTCTTATGGATCCAAAC	2005
Qy	585	LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn	604
Db	2006	CTAGGAAGGCTGTCTGTTATGTCCAATTCCTCAGAGGTTTGCAGGTATCAGTAAAT	2065
Qy	605	AspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGly	624
Db	2066	GATCGATATGCAACACAGAAACCGTGTTCATATTAATTTGAGAGGTCTCGATGGT	2125
Qy	625	IleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGly	644
Db	2126	CTCCAAGACCAAGTTTATGTGGGAACCGTTGTGTTCACAGACAGCTCTTTATGGT	2185
Qy	645	TyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLys	664
Db	2186	TATGAACCCCCCAATTAAGCAGAAAGAGCCAGGTACTTCTCTCGCTTTTGGGGGACGA	2245
Qy	665	LysLysAlaSerLysSerLysLysArgSerSerAspLysLysSerAsnLysHisVal	684
Db	2246	AAGNAGCAAAAAGTCCGAGGAGAGACCGGAAAAAGAAAAGTCAACAAACATGTG	2305

Percent Similarity:	93.99%	Conservative:	69
Best Local Similarity:	87.60%	Mismatches:	59
Query Match:	Indels:	6	
DB:	Gaps:	6	
US-09-900-237A-30 (1-1080) x AF200529 (1-3676)			
QY	1	MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln	20
Db	254	ATGACGGC--GCGCAGCCACGCAATTCGGGAAGCATGTGGCGGCGAGGTGTGCCAG	310
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
Db	311	ATCTGCGGCACGGCGTGGACCGCGCGGAGCGCGACCTTTCACGGCTCGCAGCTC	370
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla	60
Db	371	TGGGGCTTCGCCGTGCGGCCCATGCTACAGTACAGCGCAGGACGCCACCGAGCG	430
QY	61	CysLeuGlnCysLysThrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
Db	431	TGCCCCAGTGCAGAGTAAGTACAAGCGCCACAAGGGAGGCCACCCAGTACACGGTGAG	490
QY	81	GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
Db	491	GAATAATGAGATGTGGATGCTGACGATGTGAGTGACTACAACCTACCAAGCATCTGGCAAC	550
QY	101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly	120
Db	551	CAGGATCAGAAGCAAAAGATTGCTGAGAGATGCTCACTTGGCGGCAAACTCACGTGGC	610
QY	121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
Db	611	AGT---GATATTGGCTCGCTAAGTATGACAGCGGTGAAATTTGGGCATCGGAAGTATGAC	667
QY	141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160
Db	668	AGTGTGAGATCCCTCGTGGATATATCCGTCACCTAACTCATAGCCAGATCTCAGGAGAG	727
QY	161	IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg	180
Db	728	ATTCCTGGAGCTTCCCTCGAT---CATATGATGTCTCCTGTTGGGAAATTTGGCAGGCGT	784
QY	181	Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer	199
Db	785	GGACATCAATTTCTTATGTAAATCATTTCTCAAAACCCATCGAGGAGTTCTCCGGTAGC	844
QY	200	IleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLysGlnAspLysGly	219
Db	845	CTTGGCAATGTTCATGGAAGAGAGGGTGGATGGATGGAAATGAAG---GATAAAGGT	901
QY	220	AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp	239
Db	902	GCAATTCCTATGACCAATGGAACAAGCATTTCTCCATCAGAAGGGCGTGGAGTTCGTGAT	961
QY	240	IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln	259
Db	962	ATTGATGCTCTACTGATTAATTAACATGGAAGATCCCTTACTGAAATGATGAACTCGGCAA	1021
QY	260	ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle	279
Db	1022	CCTCTATCTAGAAAAGTGCCAAATTCCTTCATCCAGAATAAATCCGTACAGAATGGTCATT	1081
QY	280	ValIleArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg	299
Db	1082	GTGCTACGTTGGCTGTCTATGCATATTTCTGGCTACCGTATCCATCACAATCTGTGAC	1141
QY	300	AsnAlaTyrProLeuThrLeuLeuSerValIleCysGluIleThrPheAlaLeuSerTyr	319
Db	1142	AATGCATATCCATCTGGCTTTTATCCGTATATGAGATCTGGTTGCTTGTCTCTGG	1201
QY	320	IleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeu	339

Db	1202	ATTTTGGATCAGTTCCTCCAAAGTGTGCCCAATCAACCGGTGAACAATACCTTGTATAGACTG	1261
QY	340	AlaLeuArgTyrAspArgGlyGluProSerGlnLeuAlaAlaValAspIlePheVal	359
Db	1262	GCITTAAGGTATGACCGAGAGGTGAACCATCTCAATTAGTCTCTGTTGATATTTTGTGTC	1321
QY	360	SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle	379
Db	1322	AGTACTGTGGATCCATTAAGAGGAGCTCTCTGTCTCACTGCAATATCTGTGCTTCCATC	1381
QY	380	LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer	399
Db	1382	CTTGCTGTCAATATCGGTTGCAAGGTATCTTGCTATGTTTGGATGATGGAGCTGCT	1441
QY	400	MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPhe	419
Db	1442	ATGCTGACTTTTGATGCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGTTCGGTTC	1501
QY	420	ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp	439
Db	1502	TGTAAGAGGTACAAACATAGAGCCCTANGGCCCGGAAATGGTACTTTGCTCAGAAAATTGAT	1561
QY	440	TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysValProGlu	459
Db	1562	TACTTGAAGACAAAGTTCAAACCTCATTTGTGAAGAAGCGCGGCCCATGAAGAGAGAA	1621
QY	460	TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu	479
Db	1622	TATGAAGAATTCAAAGTTCTGATCAATGCTCTGTAGCCCAAGCCACAAAAGTTTCCGAG	1681
QY	480	GluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisPro	499
Db	1682	GAGGATGATCATGCAAGATGCTACCTTGCTGGCTGGGAAACAATACTAGGACCATCT	1741
QY	500	GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu	519
Db	1742	GGAAATGATCAGTWTTCCTGGGTCAAGTGGAGGGCTTGACGTTGAAGGCAATGAACAT	1801
QY	520	ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla	539
Db	1802	CCTCGTTTGGTTTATGTCTCGTGAAAAAGCTCTGGATTCACACATCACAGAAGGCT	1861
QY	540	GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu	559
Db	1862	GGTGCCATGAATGACCTTTGTTGCTGTATCAGCTGTCTTACTAATGGGCAATACATGTTG	1921
QY	560	AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe	579
Db	1922	AATCTTGATTGTGACCACTACATCAATAATAGCAAGGCTCTTCGAGAAGCTATGTGCTTC	1981
QY	580	LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp	599
Db	1982	CTTATGACCCCAACCTTAGGAAGGAATGTCTGTTATGTCCAATTTCTCAGAGGTTTGTAT	2041
QY	600	GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu	619
Db	2042	GGTATTGATAGAAATGCCGATATGCAAAACAGGAACACCTGTGTTTTTCGATATTAACAT	2101
QY	620	ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg	639
Db	2102	AGAGGTCTTGACGGCATTCGAGGCCCATTTATGTGGGAACTGTGTTGTGTTTAAACAGA	2161
QY	640	ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer	659
Db	2162	ACGGCTTTATGTGTTATGAGCTCCAGTCAAGAAAAAAGCAGGCTCTTCTCTCTTCG	2221
QY	660	LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys	679
Db	2222	CTTTGTGGGGAGAGAAAAGACGTCAAAATCTAAGAAG---AGCTCGGAAAAAGAAAG	2278
QY	680	SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyGlu	699
Db	2279	TACATAGACAGCAGACAGTCTCTGTACCAGTATTAACTCGAAGATATAGAGGAGGG	2338

FEATURES
SOU


```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="OJ1559_F09"
complement(2210..2833)
/gene="OJ1559_F09.101"
complement(join(<2210..2425,2690..>2833))
/gene="OJ1559_F09.101"
/notes="start and end point are not identified"
complement(join(2210..2425,2690..2833))
/gene="OJ1559_F09.101"
/notes="predicted by FGENESH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC57272.1"
/db_xref="GI:28411797"
/translation="MAADAPMLGTMVTVEETPYDLDPPTSKLADAHQGTTRPCR
TKDVLPSITFAWVITVCLVHDDLSLIQTLLFELVWVVIQWLIVFMTYSAPDDAN
QEYRKRYNVVNAWFC"
complement(join(4610..4836,5286..5394))
/gene="OJ1559_F09.102"
complement(join(4610..4836,5286..5394))
/gene="OJ1559_F09.102"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(join(6618..6632,7157..7330))
/gene="OJ1559_F09.103"
complement(join(6618..6632,7157..7330))
/gene="OJ1559_F09.103"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(7690..9417)
/gene="OJ1559_F09.104"
complement(join(<7690..7908,7998..8513,8614..>8634,
8736..8906,8992..9123,9208..>9417))
/gene="OJ1559_F09.104"
/notes="start and end point are not identified"
complement(join(7690..7908,7998..8513,8614..8634,
8736..8906,8992..9123,9208..9417))
/gene="OJ1559_F09.104"
/codon_start=1
/product="putative polygalacturonase Pg1"
/protein_id="BAC57273.1"
/db_xref="GI:28411798"
/translation="MACNVLIISAAYVVMVSFLLCSVHAKVNNASFHILTSQDSDLRGR
SLESKVDFVRKHGAYGGQDDTKALAKAAACSSSQPSIVILIPKGRYLTKHIT
LSGPCSSITFMIEGTLVAPKRSQMSKETIRHWMFNGVSLTVAGGTVDENGKIW
WQNSCKTNALPCTEPTALPTFYSNLKVNLLKLLNSQIIMSVEDCTDVRISLTI
TAPGTSPNTDGIHITRSKNVQVTGCIITGDGCMSEIDGTENLHVKNVCGPGHGISI
GSLGDHSEAHVNVTVTRLYGTNGARKITWQGGHGYAKNIIVFQNMWENWNPPI
LIDQNVCDSPATCKEQTSAVQVSNVVFNRKIRGTSAKKAIKLDCSRNVPCQGITLKD
V
KLTIKGSGSDAKSTCGNAKWKSGIALCFQ"
join(10069..10162,10373..10608,11858..12025)
/gene="OJ1559_F09.105"
join(10069..10162,10373..10608,11858..12025)
/gene="OJ1559_F09.105"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
12214..117342
/gene="OJ1559_F09.106"
join(12214..12339,12451..12516,13288..13553,14328..14452,
14845..14938,15622..16079,16691..17342)
/gene="OJ1559_F09.106"
/notes="supported by full-length cDNA(s): AK109769"
join(13400..13553,14328..14452,14845..14938,15622..16079,
16691..16966)
/gene="OJ1559_F09.106"
/notes="contains full-length cDNA(s): AK109769"

```

```

/codon_start=1
/product="NLI interacting factor-like protein"
/protein_id="BAC83123.2"
/db_xref="GI:38175433"
/translation="MOTRKKGATKNALGDPASTTKTSRQPREAAQAAASEKKVNDLITS
SAKKKSVGAPSKKNRASKGRKLIISACDAANSENVSVQVGGIPHOQSDNDVGRP
CNSLFPAYHLOKECGASNFPAKLEHGRGDTLGSVSSVEERTTHAQGRKVTSSRST
SHAVKTCVGDHHTLNAQSAFCNTPLEBEDEFSELGNSLSEVSAIYLAQVQSKLCIDE
HSQDSITGEVDPDEEDDFDYAFIKDLPDLVLPKFRPVLPLPKQTRSCPITTT
LKHLICQLNHARMLTSHFOILFTSENTQMSDAVPIKSKSWRGLPICLRQSFSLQVK
AFMQNSFMFLILKESCVIVFTVPIVPMRETT"
complement(19106..20743)
/gene="OJ1559_F09.107"
complement(join(<19106..19339,19456..19971,20079..>20270,
20360..20491,20600..>20743))
/gene="OJ1559_F09.107"
/notes="start and end point are not identified"
complement(join(19106..19339,19456..19971,20079..20270,
20360..20491,20600..20743))
/gene="OJ1559_F09.107"
/codon_start=1
/product="putative polygalacturonase precursor"
/protein_id="BAC57275.1"
/db_xref="GI:28411800"
/translation="MASTVLIISASLLALFFFLHSADADVGSNVFSIQSYGAHGDGRH
DOTKALGDTWAACSSAKPAVLIPKGRKYLTKHITLSGPKSSISLMVKGSLVASPE
RSDMSKETIRHMLISGVTGLTVGGTIDNGKIKWQNSCKTNKLPCTEPTALTFT
YSCNLIKBYLVKVNQQIQISVEDCTDVMVSRLSITAPETADPTDGIHITRGRDVEV
TDCMIKTGDDCMSIEDGTENLHVKNVCGPGHGISISGLGDHSEAHVNVTVNDVRL
YGTANGARIKTWQGGKSAKNIIVFQNMVNDVNPILIQNYCDSTPCPKQKSAVEV
SNVLPKNIRGTSAEAIMHCSSVPCGHITLNVNLIVKGGSSNAKSTCQNAEWKK
SGSVSPQPCGFRN"
24140..25351
/gene="OJ1559_F09.108"
<24140..>25351
/gene="OJ1559_F09.108"
/notes="start and end point are not identified"
24140..25351
/gene="OJ1559_F09.108"
/notes="similar to Oryza sativa chromosome7, P0418E08.23"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC57276.1"
/db_xref="GI:28411801"
/translation="WAPPRKRRRVDSGGNGNYPLVATPNDGVLVLDLNAVLRL
PASPACRLRAVCRPWRVLSDFRFAAHAARHPDPLVVAACDLDDAGGIELVDVLYV
GASGDVAKRVFAGRCDDTDAVSARDGVALLVGNDRRLVLDAAASAAVSLVPDGEHP
INCSFTLGRAASSSGEHKVLRTGVVHGEPQCAVLTLAVAGRGQNAWRAPSPFL
VVRTRRGDVAVAGVAVYFLLRAYLADRIAPDLEAQWRPALVGPFLPAWRPFRPD
RPRVTLAELGSLVLAIDHRAATLIDMLFLAAGDGEQHWKQYTTVTWPHRRPWRCD
GERAPVVLDDGRIVFWWAGGSGGTHRGGVWVRVDPITGGHTDVAARCAHVGV
YTGNLLSLVSE"
complement(join(26590..26940,27083..27243,27365..27502,
28511..28604))
/gene="OJ1559_F09.109"
complement(join(26590..26940,27083..27243,27365..27502,
28511..28604))
/gene="OJ1559_F09.109"
/notes="supported by full-length cDNA(s): AK062943"
complement(join(26842..26940,27083..27243,27365..27497))
/gene="OJ1559_F09.109"
/notes="contains EST(s): AU089698(C2176),D23094(C2176)
contains full-length cDNA(s): AK062943"
/codon_start=1

```

Alignment Scores:

Pred. No.:	0	Length:	87792
Score:	4838.00	Matches:	991
Percent Similarity:	62.37%	Conservative:	50
Best Local Similarity:	59.38%	Mismatches:	37
Query Match:	83.73%	Indels:	596
DB:	8	Gaps:	14


```
US-09-900-237A-30 (1-1080) x AP003837 (1-87792)
QY      3 GlyAspAlaasp---AlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIle 21
Db      64877 GGTGATCTATCTGTTGGCAGAGAGTCCGGAGGACACGGAGCGGCGAGGCGTCCAGATC 64818
QY      22 CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
Db      64817 TCGCGGACGGCGTGGGACACGACGGGCGGAGGCGTGTTCGCCGCTCGACGCTGCG 64758
QY      42 ArgPheProValCysArgProCysTyrGluHisGluArgGlyGlyThrGlnAlaCys 61
Db      64757 GGCTTCCCGGTGCGCGCCCTGCTAGAGTACGAGCGGAGGATGGCACCCAGGCTTGC 64698
QY      62 LeuGlnCysLysThrLysTyrLysArgHis-Arg- 72
Db      64697 CCCAGTGCAGACCAAGTACAGAGGCCACAGGGTAAATGCTACCATCTCTCTCTCTC 64638
QY      72 ----- 72
Db      64637 CTCTCCCTCCTCTCCATTTCCGCGCATCTCCGCGCCATCGACTGTAAATAAGTTCAGAAA 64578
QY      72 ----- 72
Db      64577 CATCTCAGAAAAAAGTCTGGATCTCTAGTTCAGAACAGTCAAAACGAGCATTTCTTAT 64518
QY      72 ----- 72
Db      64517 AGATGGATCCTGTGCTTCTTGTAGGAAGGTCATACAGATCCAATGTTTGTGATTT 64458
QY      72 ----- 72
Db      64457 TTTGGAAATTCAGAAAAAATAATGTCGATTAGTAAACGATGTTACAGTCAGTCAAGGC 64398
QY      73 -----GlySerProAlaIleArgGlyGluGlyAs 83
Db      64397 TTTGCTGAACGTGTATGGAAATTAATCAGGAGGCCCGCGATCCGTGGGAGGAAGGCGA 64338
QY      83 pAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspG 103
Db      64337 GGATACTGATGCTGATGATGTCAGTGACTACAACTACCTGCTGTCATCGCAGTGGCGACCA 64278
QY      103 nLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySerGlyAs 123
Db      64277 GAAGCAGAAGATTGCTGATAGGATCGCAGTTGGCGCATGAAATGCTGGGGGGTGGGAGA 64218
QY      123 nValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyG 143
Db      64217 CGTCGGCCGTCCTCCAAAGTATGACAGTGGCGAGATCGGCTCACCAAGTATGACAGTGGCGA 64158
QY      143 uIleProArgGlyTyrValProSerValThrAsnSerGlnMet----- 157
Db      64157 GATCCTCGGGATACATCCCTTCAGTCACTAATAGCCAGGT--GGGTGCGAGTTTTTAAC 64099
QY      157 ----- 157
Db      64098 AGATCGTGATATCAGTTGTCGGGTATGCTCAAAATGATTAAATCTTGAAATGGAATTTGCA 64039
QY      158 ---SerGlyGluIleProGlyValaSerProAspHisHisMetMetSerProThrGlyAs 176
Db      64038 GATCTCGGAGAGAAATCCCTTGGAGCTTCCCTGATCATCATATGATGCTCTCTACCGGAAA 63979
QY      176 nIleSerArgArgAlaProPheProThrValAsnHis----- 188
Db      63978 CATTCGCAAGCGTCTCCATCCCTATATGTGAACCATTCACGATGCTTTATTTCTTGATA 63919
QY      188 ----- 188
Db      63918 TGATGTCACCTCTCATTTAGGTACATCCGCTTGAGATCTAAGATCGTAATPATCTTTTCTT 63859
QY      189 -----Ser-ProAsnProSerArgGluPheSerGlySerIleGlyAsnValAlaTrpL 206
Db      63858 GCTTCTTCTAGCAAAATCCATCAAGGGAGTTCTCTGGCAGCATTTGGAATGTTGCCTGGA 63799
```

382	Val	Asp	Tyr	Pro	Val	Asp	Val	Ser	Cys	Tyr	Val	Ser	Asp	Asp	Gly	Ala	Ser	Met	Leu	401
62718	GTT	GAT	TAC	CCAGT	GGAC	AGG	CTC	TCT	GTG	TAT	GTAT	CTG	ACAT	TGGT	CTG	CTCA	TGCTG	62659		
402	Thr	Phe	Asp	Ala	Leu	Ala	Glu	Thr	Ser	Glu	Phe	Ala	Arg	Lys	Tyr	Val	Pro	Phe	Val	421
62658	ACT	TTT	GAT	GCAT	TGG	CTG	AGC	CTT	CAG	AGT	TTC	TAG	AAAGT	GGG	TCC	ATT	CGT	TAAG	62599	
422	Lys	Tyr	Asp	Ile	Glu	Pro	Arg	Ala	Pro	Glu	Phe	Tyr	Phe	Cys	Gln	Lys	Ile	Asp	Tyr	441
62598	AAG	TAT	AAC	ATT	GAG	CTC	CAAG	CTC	TCT	AGT	GTG	TACT	TCT	CC	CAG	AAA	ATC	GAT	TACT	62539
442	Lys	Asp	Lys	Val	Gln	Pro	Ser	Phe	Val	Lys	Asp	Arg	Arg	Ala	Met	Lys	-----	---	---	457
62538	AA	GAC	AA	AGT	CC	ACC	CTT	CAT	TTG	TAA	GAC	CGT	CGT	CC	CA	TGA	-----	---	---	62480
457	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	457
62479	TT	ATA	AA	CAC	ATT	AA	TAT	TA	TTT	TG	CTAT	CTT	CTG	CTC	CA	TGCT	TAAC	AGT	CTT	62420
458	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	62420
62419	AAAA	AAAA	AC	CTT	TAC	AG	CAG	AG	AT	AT	AG	AA	TT	CAA	GT	TAG	GA	TAA	AT	62360
472	Ser	Lys	Ala	Leu	Lys	Val	Pro	Glu	Glu	Gly	Tyr	Phe	Ile	Met	Gln	Asp	Gly	Thr	Pro	491
62359	GCT	AAG	G	CAC	AAA	AGT	CC	CG	GAG	GAG	GAT	GAT	CAT	G	CA	AGAT	GGC	CAC	ACC	62300
492	Gly	Asn	Asn	Thr	Arg	Asp	His	Pro	Gly	Met	Ile	Gln	Val	-----	-----	-----	-----	-----	-----	504
62299	GG	AA	CA	TACT	TAG	GG	CA	CT	CCT	GG	AT	GATT	CAG	GT	-----	-----	-----	-----	-----	62241
504	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	504
62240	TAC	TAT	GAC	TAT	AGT	TGT	GTG	TAG	CTT	CTT	TCT	TACT	TG	AT	GAT	TAT	TG	CTT	GTG	62181
505	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	62181
62180	ATT	GCT	CTT	CTT	CTG	ATT	TTT	CTT	TAG	GT	TTT	CTC	TGT	CAC	AGT	GGT	GGC	CTT	GAT	62121
516	Gly	Asn	Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro	Gly	Phe	Gln	535
62120	GGT	AAT	CAG	CTT	CC	CG	CTG	CT	CAC	GT	TAT	CT	CGT	GAG	AGC	GT	CCT	GGT	TT	62061
536	His	Lys	Lys	Ala	Gly	Ala	Met	Asn	Ala	Leu	-----	-----	-----	-----	-----	-----	-----	-----	-----	545
62060	CAC	AGA	AGG	CTG	TG	CG	CA	TGA	AT	GT	CTCT	-----	-----	-----	-----	-----	-----	-----	-----	62002
545	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	545
62001	ACT	AT	CTAG	TTC	ATT	TTT	CAC	TTCT	CTCT	CTA	AA	GT	TCT	AA	TG	AGG	AT	TAA	TCT	61942
546	---	Val	Arg	Val	Ser	Ala	Val	Leu	Thr	Asn	Gly	Gln	Tyr	Met	Leu	Asn	Leu	Asp	Cys	564
61941	CAG	TT	CGT	GT	CT	CAG	CT	GT	CT	TAC	CA	AT	GG	CAC	AGT	CAT	CT	TG	AA	61882
565	His	Tyr	Ile	Asn	Asn	Ser	Lys	Ala	Val	Arg	Glu	Ala	Met	Cys	Phe	Leu	Met	Asp		

Db	61642	CTAATAATTTTTTGATGCACATAATTTTAAATTTCTAACAAATTTCAATATTTTTTTTTGGGA	61583
QY	618	-----AsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr	630
Db	61582	TATTTCTGTCTTCTACAGATTAACTTGAGGGGTCTTGATGGCATCTCCAGGACCAAGTTTAT	61523
QY	631	ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLys	650
Db	61522	GTGGAACTGGTGTGTATTCACAGAACTGCTCTATATGTTAATGAACCCCAATTAAAG	61463
QY	651	AlaIlysIysProGly-----PheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLys	669
Db	61462	CAGAAGAAGAGGGAAGTTTCTTGTCATCACTATGTGGGGGAGCAAGAGCAAGCAAG	61403
QY	670	SerLysIysArgSerSerAspLysLysSerAsnLysHisValAspSerSerValPro	689
Db	61402	TCAAGAGNAGAGTTCGGACAGAGAGAGTCAACCAAGCAGCGTGGACAGTGTGTGCCA	61343
QY	690	ValPheAsnLeuGluAspIleGluGlyValGlu	701
Db	61342	GTTTTCAATCTTCAAGATATAGAGGAGGGTGTGAAGGTACCATCCATATGTTCCAGTTA	61283
QY	701	-----	701
Db	61282	TTGCATAATTTAGGATTAAGAATTGAATCAACATATATGCATCTGCATATCATCTAT	61223
QY	702	-----GlyAlaGlyPheAspGluLysSerValLeuMetSerGlnMetSerLeuGlu	719
Db	61222	ATGGCAGGTGCTGGATTTCGATGATGAGAATCACTTCTTATGTCTCAATGAGCTTGGAG	61163
QY	720	LysArgPheGlyGlnSerAlaPheAlaAlaSerThrLeuMetGluTyrGlyGlyVal	739
Db	61162	AAGAGATTGGGCAAGTCTGCAGCGTTGTGTGCTCCACTCTCATGAAATATGTTGGTGT	61103
QY	740	ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly	759
Db	61102	CCTCAATCTGCAACCCCAAGATCTCTTTTGAAGAAGCTATCCATGTGATAAGTTCTGGC	61043
QY	760	TyrGluAspLysSerGluTrpGlyThrGlu	769
Db	61042	TATGAGGACAAGACCGAATGGGAACTGAGGTAAACATGGGCTTTCTCAGACGACATGG	60983
QY	769	-----	769
Db	60982	TTACATCTTTTTTAGTAACCTTGTTAAATTTTACACACAGGCTATGCAACTTGCATGGTA	60923
QY	770	-----IleGlyTrpIleTyrGlySerValThrGluAspIle	781
Db	60922	TGTTAATGCTGACATATGTTGCAGATGGTGGATCTACGGTTCGGTGACAGAAATAT	60863
QY	782	LeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLysArg	801
Db	60862	CTCACTGGATTCAAGATGCATGCGGAGCTGGAGATCAATCTACTGTCATGCCAAGCGC	60803
QY	802	ProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg	821
Db	60802	CCAGCTTTCAAGGGGTCTGCTCTTATCAATCTTTTCAGATCGTCTTAACCAAGTGTTCGG	60743
QY	822	TrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTrpGlyTyr	841
Db	60742	TGGGCACCTTGGTCTGTGTGAATCTTTTTCAGTGGCCATGTGCCATATGGTACGGCTAT	60683
QY	842	GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeu	861
Db	60682	GGAGGACGCTTAAAGTTCTTGGAGAGATTGGCTTACATCAACACCATATTATCCATTG	60623
QY	862	ThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLys	881
Db	60622	ACATCATCCGCTTCTCATATACGTGTGTGTGCTGTCTATCTGTGCTCACCTGGGAAG	60563
QY	882	PheIleMetProGluIleSerAsn	889

Db	60562	TTTCATATCCAGAGTAAGCAATGACTGCTTATACATACCTTCTATAGAAAATGTTTG	60503
Qy	890	-----Leu-AlaSerIleTr	894
Db	60502	AGCAACCAAACTAACTAAATTAATTTTGGAAATGCAGATTAGCAACTTTGCTAGTATTGG	60443
Qy	894	pHeIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGl	914
Db	60442	GTTCATCTCTCTCTTCATTTCAATTTTGGCACTGGTATCTCTTGAGATGAGGTGGAGTGG	60383
Qy	914	yValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAl	934
Db	60382	TGTTGGATCGATGAGTGGGGAGGAATGAACAGTTCTGGGTTATTGGAGTATATCTGC	60323
Qy	934	aHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPh	954
Db	60322	GCATCTTTTGGCGTCTTCAGGGTCTCTCAAGGTGTTGCTGGTATCGACCACTTT	60263
Qy	954	eThrValThrSerIysAlaAsnAspGluGluGlyAspPheAlaGluLeuTrpMetPheLy	974
Db	60262	CACGTGCACCTCAAGGCTTCTGATGAAGATGGCGACTTTGCTGAGCTCTACATGTTCAA	60203
Qy	974	sTrpThrIleuLeuLeuProProThrIleLeuIleIleAsnMetValGlyValVa	994
Db	60202	GTGGACAACGCTCTCATCCACCAGCAGCACTTGTATCATTAACCTGCTGGTGTGTGT	60143
Qy	994	lAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLy	1014
Db	60142	TGCTGGTATCTCATACGGATCAACAGCGGTACCAATCATCTGGGACCGCTCTTTGGCAA	60083
Qy	1014	sLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheIleuLysGlyLeuMetGl	1034
Db	60082	GCTCTTCTTGGCTTCTGGGTGATTTGCCACTTGTACCCCTTCTCAAGGGTCTTATGGG	60023
Qy	1034	yArgGluAsnArgThrProThrIleValIleValTrpAlaValLeuAlaSerIlePh	1054
Db	60022	TCGGCAAAACCGCACTCCGACCATCTGTTGTTGTTGGCAATCCTCTGGCTTCGATCTT	59963
Qy	1054	eSerLeuLeuTrpValArgValAspProPheThrArgLeuAlaGlyProAsnIleGl	1074
Db	59962	CTCATTTGTTGGTTCGCATCGATCCATTACCACCGTGTCCACCGGCCAGATACCCA	59903
Qy	1074	nThrCysGlyIleAsnCys 1080	
Db	59902	AACATGTGGCATCACTGC 59884	
RESULT 11			
AC145384/c			
LOCUS	41693 bp	DNA	linear
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 3 clone		HTG 16-OCT-2003
ACCESSION	OSJNBa0060M17, *** SEQUENCING IN PROGRESS ***.		
VERSION	AC145384.2	GI:37693577	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
	Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1 (bases 1 to 41693)		
AUTHORS	Buell, R., Hsiao, J., Zismann, V., Moffat, K.M., Hill, J.,		
	Gansberger, K., Burgess, S., Jarrahi, B., Shvartsbeyn, M., Brenner, M.,		
	Ciecko, A., Pai, G., Vanaken, S., Hansen, C., Uterbach, T.,		
	Feldblum, T., Khalak, H.G., Yuan, Q., Quackenbush, J., White, O.,		
	Salzberg, S. and Fraser, C.		
	Oryza sativa ssp. japonica cv. Nipponbare OSJNBa0060M17 BAC genomic		
	sequence		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 41693)		
AUTHORS	Buell, R.		
DIRECT SUBMISSION			
TITLE	The Institute for Genomic Research, 9712		
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA		
REFERENCE	3 (bases 1 to 41693)		
AUTHORS	Buell, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-OCT-2003) The Institute for Genomic Research, 9712		
	Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 1 contigs. Gaps between the contigs		
	* are represented as runs of N. The order of the pieces		
	* is believed to be correct as given, however the sizes		
	* of the gaps between them are based on estimates that have		
	* provided by the submittor.		
	* This sequence will be replaced		
	* by the finished sequence as soon as it is available and		
	* the accession number will be preserved.		
	* 41693: contig of 41693 bp in length.		
FEATURES	1	Location/Qualifiers	
source	1..41693		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/cultivar="Nipponbare"		
	/sub_species="japonica"		
	/db_xref="taxon:39947"		
	/chromosome="3"		
	/clone="OSJNBa0060M17"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	41693
Score:	4632.00	Matches:	949
Percent Similarity:	60.06%	Conservative:	66
Best Local Similarity:	56.15%	Mismatches:	55
Query Match:	80.17%	Indels:	621
DB:	2	Gaps:	16
US-09-900-237A-30 (1-1080) x AC145384 (1-41693)			
Qy	7	AlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGlyLeu	26
Db	11151	CGGAGAAATCGGGGAACAA-----TGCCATGCTGCCAGATCTGCGGCGAGCGGTG	11098
Qy	27	GlyThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProValCys	46
Db	11097	GGCAGCGGGGACGGCGAGCTCTTACCGCTCGGAGCTGCGGGTCCCGGTGTC	11038
Qy	47	ArgProCysTyrGluHisGluArgGlyThrGlnAlaCysLeuGlnCysLysThr	66
Db	11037	CGGCCATGCTACGAGTACGAGCGCAAGGACGCGGAGTCCCGGAGTGCAGGACC	10978
Qy	67	LysTyrLysArgHisArg-----	72
Db	10977	AAGTACAGCGCCACAA--GGGTATAAGTTGCACATTCATTCATCTCTCTTGGATT	10919
Qy	72	-----	72
Db	10918	TCACCTAAATACTTCATCTTTGGTTAAGAACAGCTACATTCGATCGAATCTCGTGAC	10859
Qy	72	-----	72
Db	10858	CCAAATCAGGTTTGTGCATATTCATTTAGCAATAATAGTAGAAGATATATATATGCT	10799
Qy	72	-----	72
Db	10798	TAGTGAATATGCTGAGTTAATAGCACAGAGTCTCTTCTTTAGTGAGCTGTATTGGAAGT	10739
Qy	72	-----	72
Db	10738	TGGTAGTGACAGTGACAACTAATTTCTTCTCTGTATGGGTATTTTCCATTGAGAGAAG	10679
Qy	73	-----	75
Db	10678	CAAAATGTAGTGAATTTGTTGAGCTTAACAAGATATTTAATCTTTGCTCAGGAGCCCG	10619

QY 76 AlaIleArgGlyGluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyr 95
Db 10618 CCGATCTTGGGATGAAGCGATGATTTGATCGGATGATGCTGATGTGAACCTAT 10559
QY 96 ProAlaSerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArg 115
Db 10558 CCAACATCCGGCAACAGGACCAACACACAGATTCGCGAGAGATGCTCACCTCGCGC 10499
QY 116 MetAsnThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGly 135
Db 10498 ATGAATCTCTGGAGGAATGATATATTGTCATCTCAAGTATGATATGCTGGATGGT 10439
QY 136 LeuSerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSer 155
Db 10438 CATCCCAAGATGACAGATGTTGAATCCCTCGCATATATATCCATCGCTCACTCACAGC 10379
QY 156 Gln----- 156
Db 10378 CAGGTTTGTAGAAATTTTGTTCGATAGCGTGAAGTCGTTAACCTTCAGGGTATGTGCAA 10319
QY 157 -----MetSerGlyGluIle 161
Db 10318 AGATTGAACCTAATATTAGTTTCTCTCTCGATTAAATCTATAGATCTCAGGTGAAT 10259
QY 162 ProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgAla 181
Db 10258 CCTGGAGCTCCCTCAT---CATATGATGTCTCCCGTTGGGAACATTGGCAGACGTGG 10202
QY 182 ---ProPheProTyrValAsnHisSer----- 189
Db 10201 CATCCATTTCCCTATGTGAACCATTCACGTAAAGTTTATCTATCATCAAGATCTCATGTTT 10142
QY 190 -----ProAsnProSerArg 194
Db 10141 CGTTCATCGACTGGGAAATTTAATTTATCTTTGTTGCTCAACAGCAACCATCAAG 10082
QY 195 GluPheSerGlySerIleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMet 214
Db 10081 GAGTTTCTGGTAGCTTGGCAATGTTGCATGAAAGAGAGATAGATGGCTGGAAATG 10022
QY 215 LysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGly 234
Db 10021 AAG---GATAAAGGTGCAATTTCCCATGGCTAATGTTACTAGCATTTGCTCTTCAAGAGT 9965
QY 235 ArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeu--- 253
Db 9964 CGTGGGTTGGTGACATTTGATGCATCTACTGACTATAACATGGAGGACGCCCTTACTGTGA 9905
QY 253 ----- 253
Db 9904 GTTCTGATATGCTTTTGTCCACCGTTAGTCATCTATGAATTTCTACCAGTAATTAATTC 9845
QY 253 ----- 253
Db 9844 CTTGTATCTGGTGACTAACTTGCATAGTTTGTACTTGGAGAACATAGGAAAATTC 9785
QY 253 ----- 253
Db 9784 TGCATAATTCGAATATAATTAATTTTGTGTTTACAAGTATACCTTGTCTGTAACGAGGC 9725
QY 253 ----- 253
Db 9724 AATATGCACATTTAATGATCAGATTTAAATGGTTGTGATCTCTGTAACCTTATAATGC 9665
QY 254 -----AsnAspG 256
Db 9664 GGACTTATCATATAGCTTGGCACCATAAAATGTTGCTATCTGTTATACAGGAATGATGA 9605
QY 256 uThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrAr 276
Db 9604 AACAGCCAACTCTATCGAGAAAAGTGCCAAATTTGCTCATCCAGAAATAATCCGTACAG 9545
QY 276 gMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAs 296

Db 9544 GATGGTTATTGTGCTCCGATGATGTCTCTATGTATATTCTTGCACTACCGTATCACAAA 9485
QY 296 nProValArgAsnAlaTyrProLeuTyrPheProLeuLeuSerValIleCysGluIleTyrPheAl 316
Db 9484 TCCTGTGCGTAATGTCATACCCGCTATGTTGCTCTCTGTCTATATGTGAGATTGTTGTC 9425
QY 316 aLeuSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLe 336
Db 9424 TCATATCTGATTTCTGGATCAGTTCCTCCAAAGTGTCCCAATCAACCGTGAAACTTACCT 9365
QY 336 uAspArgLeuAla----- 340
Db 9364 GGATAGCTGGCTTTAAGGTGATGACAAATATTACCTCTGATACACCATTCCTCGCGTTA 9305
QY 341 -----LeuArgTyrAspAr 345
Db 9304 GCATCTAGTACTGATTAAGTACATAAATAATATCACTTGCACCTCTCAGGTATGACCG 9245
QY 345 gGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLe 365
Db 9244 GGAAGGTGAACCATCTCAATTTGGCTCTCTGTCATTTTGTCACTGCTGTGATCTCTAT 9185
QY 365 uLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPr 385
Db 9184 GAAAGAACCTCTCTGTCTCACTGCCAATACTGTGCTCTCATCTCTGTGTGATTTACCC 9125
QY 385 oValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAl 405
Db 9124 TGTGTATAAGGTATCTTGTCTGATGTCTGATGAGCTGCAATCTTAACATTTGATGC 9065
QY 405 aLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheValLysLysTyrAspIl 425
Db 9064 GTTGTCTGACCTTCAGAAATTTGTAGAAAATGGGTACCGTTCTGTAAAAAGTACAGCAT 9005
QY 425 eGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysVa 445
Db 9004 AGAACCCAGAGCTCCAGAGTGTACTTTCACACAGAAAATTTGATTACTTGAAGACAAAAT 8945
QY 445 IqlnProSerPheValLysAspArgAlaMetLys----- 457
Db 8944 TCAGGCTTCTTTTGTAAAGATCGTCTGCCATGAAGTAAATTTGGTTTATATGTGTCT 8885
QY 457 ----- 457
Db 8884 ACTTTTTTGTGTTTTTTTTTGTGTTAATCTCGCATTAATGCTGTGCAAAACCAAAA 8825
QY 458 -----ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuL 476
Db 8824 TCCAGAGGGAATATGAAGAAATTTAAAGTTCTGTGTTAATGCTCTCGTCGCCAAGGCACAGA 8765
QY 476 ysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrA 496
Db 8764 AAGTCTCTGAGGAAGGATGGATTATGCAAGATGSCACACCTTGGCTTGGTAAACATAACA 8705
QY 496 rgAspHisProGlyMetIle----- 502
Db 8704 GAGACCATCTGGAAATGATTCAGGTAATATCATCTTAGATATTCCTATGTTGGTGTGTT 8645
QY 503 -----GlnValPheLeuGlyHis 508
Db 8644 ATTGATCCAGCAATTTTCTTACTCTTGACGCTCTGTCATGCTCAGTTTCTTGTGTCAT 8585
QY 509 SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 528
Db 8584 AGTGGAGCCCTTGATACCGAAGGCAATGAGCTTCTCTGTTAGTCTATGTCTCTCGTGA 8525
QY 529 LysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAla-Leu----- 545
Db 8524 AAACGTCTCGGATTCACACCATAAAGAGCTGCTATGAATGCACTTGTGAGTAGC 8465
QY 545 ----- 545

8464 AGCTAGAGTGGTGTGCTTAAGTCAGAGATATCTTGTTTAATTCATTACTTGATATGTG 8405
Db
546 -----ValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAs 560
Qy
8404 GAATTTGAAACTGCAGGTTCGTGTATCAGCTGTCTTACTAATGACAATACTTGTGAA 8345
Db
560 nLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLe 580
Qy
8344 TCTTGACTGTGATCACTACATCAACATAGCAAGCTCTCGAGAGGATATGTGCTTCCT 8285
Db
580 uMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGl 600
Qy
8284 TATGATCCAAACCTAGAGGGGTGCTGTATGTCCAAATTCCTCAGAGGTTTGACGG 8225
Db
600 YIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp----- 616
Qy
8224 TATCGATAGAAATGATCGATATCAAAACAGGAACACCGTGTGTTTCGATGTAAAGCTTGGT 8165
Db
616 ----- 616
Qy
8164 TCCCTTTAAGATTCTTGGATGCTTCATCTCAGCTTTAGCAAAACATCAATAACTGTTGAT 8105
Db
616 ----- 616
Qy
8104 GGATCAGCAATATTTATGTTTCTATGTGATGGAATTTGAACATTTCTGGAGTTTGTCAAGT 8045
Db
617 -----IleAsnLeuArgGlyLeuAspGlyIleGlnGl 627
Qy
8044 AACTTATCTGCATTTCTGCTGCAGATTAATTTGAGAGGTCTCGATGGTCTCCAAGG 7985
Db
627 yProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluPr 647
Qy
7984 ACCAGTTTATGTGGAAACCGGTGTGTGTTCAACAGAACAGCTCTTTATGGTTATGAACC 7925
Db
647 oProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLeuAl 667
Qy
7924 CCCCATTAAAGCAGAGAGCCAGGTACTTCTCTCGCTTTGTGGGGACGAAAGAAC 7865
Db
667 aSerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSe 687
Qy
7864 AAAAAAGTCGAAGGAGAGAGACGCGAAAGAAAAAGTCAACAAACATGTGGACAGTTC 7805
Db
687 rValProValPheAsnLeuGluAspIleGluGlyValGlu----- 701
Qy
7804 TGTGCCAGTTTTTAATCTTGAAGATATAGAGAGGGATTGAAGTGACATCGAGTCTGG 7745
Db
701 ----- 701
Qy
7744 TTTTGTTCATATCATATGTTTTGTCTCTAGATTTTATGTATATTTGGCCACCACCAT 7685
Db
701 ----- 701
Qy
7684 GCATTTGCCACCCCAACACTTTTTTCGCAGTATTTCTGTATATGCTTAGGAAGGGAGGACA 7625
Db
702 -----G1 702
Qy
7624 CTAGGCTTAACATTTCTCTTCACTATATATTTAGCCTGGCAGCATCAATTTTGTGGCAGG 7565
Db
702 yAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPh 722
Qy
7564 TTCTGGATTTGATGAGAAATCACTACTGATGCTCAATGAGCTTAGAGAAAGATT 7505
Db
722 eGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSe 742
Qy
7504 TGTCAATCTAGTGTGTTTTGTAGCCTCCACTCTGATGAATATGGTGGTGTCTCCTCAATC 7445
Db
742 rSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAs 762
Qy
7444 CGCAACTCCAGATCTCTCTGAAGAGGCCATACATGTTATCAGCTGTGGCTATGAAGA 7385
Db
762 pLysSerGluTrpGlyThrGlu----- 769
Qy
7384 CAAAAGTACTGGGGAACCTGAGGTAAATAAGCTATGGCATTATGCTCTGCATTACATGC 7325
Db

Qy 769 ----- 769
Db 7324 TACTAGCTTACCAGTACAAATTATACAATGTACAGAAATGTGAGATATCAAGTGGTC 7265
Qy 769 ----- 769
Db 7264 TTTCTTGTACAAAATTAAAGGAGCAATTTGATTTCTGTCTGATTCGACAAAGTACTTCT 7205
Qy 769 ----- 769
Db 7204 GAGTTTCATGCGAAGAAACTTGACTAAATTGGCATCTCTCTCAGTTGCCATTATAACTTT 7145
Qy 770 -----IleGlyTyrIleTyrGlySerValThrGluAspIl 781
Db 7144 GTTTATGATGCTATGCTACTTGCAGATTGGATGATCTAUGTTCTGTTCAGAGAATAT 7085
Qy 781 eLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLysAr 801
Db 7084 TCTCACTGGATTCAGATGTCATGCGTGGTGGCGGTCAATCTACTGTCATGCCCTAAGCG 7025
Qy 801 gProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuAr 821
Db 7024 ACCAGCATTCAGGGGTCTGCCCAATTAATCTTTCCGATCGTCTGAACACGAGGTCTTCG 6965
Qy 821 gTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGlyTy 841
Db 6964 GTGGCTCTTGTTCTCTGGAAATTCCTTTAGCAGGATTTGCCATATGGTATGGGTA 6905
Qy 841 rGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLe 861
Db 6904 CGGAGGACGGCTTAAGTTCTCTGGAGAGATTTGCTTACATCAACACCACTATTATCCACT 6845
Qy 861 uThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLy 881
Db 6844 CACATCGATTCCACTCTCTTGTACTGCATATGCGAGTATCTCTCTCTCCTCCTCAGTGGAA 6785
Qy 881 sPheIleMetProGlu----- 886
Db 6784 GTTCATCATACGAGGTAAATATGTAGAAATGTTTTCTTCTTCGCAGACGAATCGC 6725
Qy 887 -----IleSerAsnLeuAlaSerIleT 894
Db 6724 AGAACAGCATCGCTGACTGACTTATTTTGGACTGCAGATTAGCAACTTTGCGAGTATT 6665
Qy 894 rPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerG 914
Db 6664 GGTATTATCTCTCTCTCTGTCATCTTTGTCTACTGGTATCCTTGAGATGAGTGGAGTG 6605
Qy 914 lYValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerA 934
Db 6604 GHTTGGCATCGATGAATGGTGGAGGAATGAGCAGTTTGGGTATTGGAGGTATTTCG 6545
Qy 934 laHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnP 954
Db 6544 CTCATCTATTTCCGGTCTTCCAGGGTCTGCTGAAGGTGCTGCTGCGAATTGATACCACT 6485
Qy 954 heThrValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheL 974
Db 6484 TCACGTCTCACCTCAAAGGCCCTCGATGAAGAAGGGGATTTCTGCTGAGCTCTACATGTTCA 6425
Qy 974 ysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyValV 994
Db 6424 AGTGGAACAACTTCTGATCCCAACGACCACTATTTTGTATCATCAACCTGGTGGTGTG 6365
Qy 994 alAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyL 1014
Db 6364 TTGCTGGTATCTCTACGCTATCAACAGTGGCTACCACTGATGGGTCCGCTCTTTGGAA 6305
Qy 1014 ysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetG 1034
Db 6304 AGCTCTCTCTTCTGCTTCCAGCTGTATGTCACCTGTACCCCTTCTCTTAAGGGTCTCATGG 6245

QY 1034 lYArgGlnAnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIleP 1054
 Db 6244 GTGGCAGAACCGTACACCGACCATTTGTTGTTTGGGCATCTCTCTTGCATCGATCT 6185
 QY 1054 heSerLeuLeuTrpValArgValAspProPheThrArgLeuAlaGlyProAsnIleG 1074
 Db 6184 TCTCTTGGCTGTGGTTCGTATCGATCATTCACACCGGTGTACAGCGCCGTATACCC 6125
 QY 1074 lNThrCysGlyIleAsnCys 1080
 Db 6124 AAAATGTGTATCACTGC 6105

RESULT 12
 AC135958 154555 bp DNA linear PLN 25-APR-2003
 LOCUS Oryza sativa chromosome 3 BAC OSJNBa0059E14 genomic sequence,
 DEFINITION complete sequence.
 ACCESSION AC135958
 VERSION AC135958.2 GI:29837774
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 154555)
 Buell,C.R., Yuan,Q., Ouyang S., Liu,J., Gansberger,K., Jones,K.M.,
 Overton II,L.L., Tsitrin,T., Kim,M.M., Bera,J.J., Jin,S.S.,
 Fadrosch,D.W., Tallon,I.J., Koo,H., Zisemann,V., Hsiao,J., Blunt,S.,
 Vanaken,S.S., Riedmuller,S.B., Utterback,T.T., Feldblyum,T.V.,
 Yang,Q.Q., Haas,B.J., Suh,B.B., Peterson,J.J., Quackenbush,J.,
 White,O., Salzberg,S.L. and Fraser,C.M.
 Oryza sativa chromosome 3 BAC OSJNBa0059E14 genomic sequence
 Unpublished
 2 (bases 1 to 154555)
 Buell,R.
 Direct Submission
 Submitted (01-NOV-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 3 (bases 1 to 154555)
 Buell,R.
 Direct Submission
 Submitted (15-APR-2003) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 4 (bases 1 to 154555)
 Buell,R.
 Direct Submission
 Submitted (25-APR-2003) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
 On Apr 15, 2003 this sequence version replaced gi:24462343.
 Address all correspondence to:rice@tigr.org

BAC clone OSJNBa0059E14 is from Oryza sativa chromosome 3
 The orientation of the sequence is from SP6 to T7 end of the BAC
 Clone.
 Genes were identified by a combination of several methods: Gene
 prediction programs including Fgenesh (<http://www.softberry.com/>),
 genscan and Genscan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and GeneSplicer
 (Mihaila Petrea and Steven Salzberg, contact mpetrea@tigr.org),
 searches of the complete sequence against a peptide database and
 the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as unknown proteins.
 Genes without protein or EST similarity, that are predicted by more
 than two gene prediction programs over most of their length are
 annotated as hypothetical proteins. Genes encoding tRNAs are
 predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
 identified by repeatmasker (Arian Smit,

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES
 source This BAC overlaps with rice BAC OSJNBb0015I02 (AC135563).
 Location/Qualifiers
 1.154555
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="3"
 /map="near C393A"
 /clone="OSJNBa0059E14"
 /note="japonica cultivar-group"
 1373..1396
 /rpt_family="AT-rich"
 complement(2338..3123)
 /gene="OSJNBa0059E14.1"
 /note="EST AU165836, AU182644 from this gene"
 complement(join<2338..2570,2665..>3123))
 /gene="OSJNBa0059E14.1"
 complement(join(2459..2570,2665..3123))
 /gene="OSJNBa0059E14.1"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AAP21417.1"
 /db_xref="GI:30103004"
 /translation="MAAIANLELSHTHTHTQVCSIDRPSRAHNGERSSQPRGGG
 GNGVCGCGGGGGVRRGAGAVRAGVRAGGAGRSVLRREGAARD
 GDDGAARAVQVGAIVRVRAPRAAPGALQARTHPRRRRHLQQDILSSGG
 2782..3013
 /rpt_family="(CGCG)n"
 4169..5246
 /gene="OSJNBa0059E14.2"
 /note="similar to ubiquitin GB:BAC06474 GI:22004048
 (Clona savigny)"
 join(<4169..4410,4495..4975,5067..>5246)
 /gene="OSJNBa0059E14.2"
 join(4169..4410,4495..4975,5067..5246)
 /gene="OSJNBa0059E14.2"
 /codon_start=1
 /product="putative ubiquitin"
 /protein_id="AAP21418.1"
 /db_xref="GI:30103005"
 /translation="WRINVTLRGDRVALDVGGATTVAQVGMVARRIAVAMORL
 FFAGCLDDHRLADYGVRRHSVFLSURLATDAYQTEHNVRLMQPTATAKQEMH
 QQQQQLHVHADDEERKAIRKPVSRALRKLILSRQYDWTQSHDAKFLDLLRHT
 GGGGARNVGELTGEDWSIRAEALNAATGSGFVEQLQRRLGELFRFEFAAGRIKNHP
 RFSYDPRRVVAKQADWKNYILENPEAAAYEGRSPRLGLRLAIFSGDGGGGGGGA
 KCRFTKARSLKRLNFRFLPKL"
 5138..5161
 /rpt_family="(CGG)n"
 9301..11079
 /gene="OSJNBa0059E14.3"
 /note="contains Pfam profile PF03763 (remorin, C terminal
 region); EST BR041106.D41920.A0907433 from this gene"
 join(9301..10065,10558..11079)
 /gene="OSJNBa0059E14.3"
 join(9403..10065,10558..10749)
 /gene="OSJNBa0059E14.3"
 /codon_start=1
 /product="remorin-like protein"
 /protein_id="AAP21420.1"
 /db_xref="GI:30103007"
 /translation="MLHSHAPPPQPEFVSQLSAPATAADVAAGDDEEVVVVTTY
 RDIHLTPPTTTTTPTRLGSRAYSWDTSRHSVSEEQFTMSREFTAMVAAGT
 TMTQPDNGNGDQLTSGIDELEETNPILAIVPDSPHPIATPARASOLEVVPVAGP
 SPAPVEARQVKKEVEVKVSAWQTAEVAKNNRFRREVINGWETQVEKASWLK
 KIERKLDSQAKALERTQNDIAKARRKAEKASAEKRLKLAKVLELANPMKAVGR
 VPTKRSFT"
 9422..9587
 /rpt_family="(CCG)n"
 complement(12645..14126)
 repeat_region
 gene

mRNA	/gene="OSJNBa0059E14.4"			Best Local Similarity: 56.15%			Matches: 55		
	/note="similar to immediate early salicylate induced glucosyltransferase GB:AAB36653 GI:1685005 (Nicotiana tabacum)"			Query Match: 80.17%			Indels: 621		
	complement(12645..>14126)			DB: 8			Gaps: 16		
	/gene="OSJNBa0059E14.4"			US-09-900-237A-30 (1-1080) x AC135958 (1-154555)					
	/codon_start=1								
CDS	/product="putative immediate-early salicylate-induced glucosyltransferase"			QY 7 AlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGlyLeu 26					
	/protein_id="AAP21423.1"			DB 16187 GCGAGAAATCGGGAGCAA-----TGCATGCTTGCAGATCTGCGCGACGCGGTG 16240					
	/db_xref="GI:30103010"			QY 27 GlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProValCys 46					
	/translation="MAAATADGCHRRRLRVFFLPFFARGHLIPMTDLACLMRAASTD AVEVEATMAYTPANAALAAATVAGNAARVVCYPPDVGLAGVECLGAAAADHTWRV YRAVDLRPAHESLHRHPDAIVADVPFWMATGVAALGVRLTFPNVGVFPPLAMN NLVARDPVEGGADPEVTPVPGMGREITIPVSELPDFLVDQDLUSMGWDRIKASQ LAGVGVVNTFAALEADYCDPFSDFARRAYFVGSPQSRAAAARRGGDGDVDCI RWLSTKQSVVYVCFGSWAHFSYQTRRELALGLEASNPFLWIRSDSGDGGRRWE PEGWERMGRGVRWAPOLAVLHPASVGFVTHCGNSVLEAAAAGVPALTWPIV FEQTNELTVEAAGFAAGVEDGGKRGVRAEATVPAGVIARAVAGFMAGGCGRR ERAAMATALESARVAVGSSWRDRLRIQLDTATASQP"			DB 16241 GGCACGCGCGAGCGCGAGCTCTTCACCGCTTCGACGCTCTCGGGTTCCTCGGTGTC 16300					
	/rpt_family="GC_rich"			QY 47 ArgProCysTyrGluHisGluAsqLysGluGlyThrGlnAlaCysLeuGlnCysLysThr 66					
repeat_region	13354			DB 16301 CGGCCATGCTACGAGTACGAGCGCAAGACGCGCAGCGGTGCCCGCAGTGAAGACC 16360					
	13913..13963			QY 67 LysTyrLysArgHisArg----- 72					
	15663..21236			DB 16361 AAGTACAGCGCCACAA -GGGTATAAGTTGCACATTATCCATTTCCTTGATT 16419					
	/rpt_family="CGG)n"			QY 72 ----- 72					
	/gene="OSJNBa0059E14.5"			DB 16420 TCACCTAAATACCTTCATCTTTGGTTAAAGACAGCTACATTCGACTGAACATCTCGTGAAC 16479					
gene	/note="highly similar to cellulose synthase catalytic subunit GB:AAC39336 GI:2827143 (Arabidopsis thaliana); EST_AU064228.AUL76398, CB096725, AA751514 from this gene"			QY 72 ----- 72					
	Join(15663..16059,16193..16379,16712..16962,17065..17164,17243..17429,17725..17991,18085..18430,18519..18656,18739..18864,18950..19162,19322..19577,19773..19975,20219..20569,20652..>21236)			DB 16480 CCAATCAGTGTGTCATATTTAGCAATTAAGTAGAGAGAGATATATATATGCT 16539					
	/gene="OSJNBa0059E14.5"			QY 72 ----- 72					
	Join(16048..16059,16193..16379,16712..16962,17065..17164,17243..17429,17725..17991,18085..18430,18519..18656,18739..18864,18950..19162,19322..19577,19773..19975,20219..20569,20652..21236)			DB 16540 TAGTGAATATCTGAGTTAATAGCACAGAGTCTCTTCTTTAGTGAGCTGTATTCGAAAT 16599					
	/gene="OSJNBa0059E14.5"			QY 72 ----- 72					
CDS	/codon_start=1			DB 16600 TGGTAGTACCAGTGACAAACAATAATTTCTTCTGTATGGGTATTTCCTATGAGAGAG 16659					
	/product="putative cellulose synthase catalytic subunit"			QY 73 ----- GlySerPro 75					
	/protein_id="AAP21426.1"			DB 16660 CAAATGTGAGTAATTTGTTGAGCTTAACAGATATTAACTTTGCTTCAGGGAGCCGG 16719					
	/db_xref="GI:30103013"			QY 76 AlaIleArgGlyGluGluGlyAspThrAspAlaAspGlySerAspPheAsnTyr 95					
	/translation="MDGAKGQCHVCQICGQGVGTAAADGBLFTACDVCQFVCRPCY EYERKDSOACQCKTKYKRHKGSPRIIGDSDSDVDADADSNVYPTSGNODHKKHIA ERMILTRMNSGRNDIIVHSKYDSGRIIGHPKYDSGEIPRIYIPSLTHSQISEIPKASP DMHGFVNIQRHGFPPVYNHSPNSPREFSGSLGNVAWKRVGDWKKRDKAIPMAN GTSIAPSEGRGVGDIIDASTDYNMDEALLINDETPLSRKVPYSSSRINPYRWIVRL IYVLCIFLHYRINTVNRNAYPLMWSYICBIFALSWILDQFPKVPINRVTYDLRAL YRDREGEPSQAPADIFVSTVDPMPKEPPLVANTVLSILAVDPVDKVSFVSDGAA MTLTDAATLSFSAKRWPFCKYSITERRAPWFAOKIDYLDKXQVASFVKRRAMK RYEFSEKRVNALVAKAQVPEGWITMODGTPNGNUTDPHGMQLQVFLGSGGLDTE GNEPLFLVTPREKRPFGPHHKGAMNALVRVSAVLNTGVQLNLDCDHYINNSKAL REAMCFDLPNIGRVYVYFFQDFGIDNRNRYANRNTVFIDNLRLDGLQGGYIV VNGCFVNRALYGEPIYQKRPFGYSSLCGRKTKKSKSEKSTKKSKSHKSDSVSP VTNLEDIEGIGSGFDDSKSLMSLEKFGQSSVEFASVTLMEYGGVQPSATPES LKAEATHVLSGVEDKSDMGTEILWYGSVTRDITLTPKWHARGWSIYCMKRRAPFK GSAPNLSDRLQALRWALGSGTEILWYGSVTRDITLTPKWHARGWSIYCMKRRAPFK GSAPNLSDRLQALRWALGSGTEILWYGSVTRDITLTPKWHARGWSIYCMKRRAPFK RNEQFWIGGISAHLFVAFQGLKVLGADIDTFTVTSKASDEEDGFADIELNFKWTLL IPTFTLLINLGVAVIGSYANSQSGWGLFGKLFAPFAPVIVHLVPLKGLMGRQN RTPITVVMWAILIASIFSLWLVRIDPFTTRVTGPDTKQCGINC"			DB 16720 CCGATCTTGGGGATGAAGCGATGATGTTGATGCGGATGATGCTAGTGTGAACTAT 16779					
gene	complement(21685..24525)			QY 96 ProAlaSerGlyThrGluAspGlnLysIleAlaAspArgMetArgSerTrpArg 115					
	/gene="OSJNBa0059E14.6"			DB 16780 CCAACATCCGCAACACGAGCACCAAGATTGCGAGAGAGATGCTCACCTGGCGC 16839					
	/note="similar to polygalacturonase precursor GB:AC26512 GI:3320462 (Cucumis melo)"			QY 116 MetAsnThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGly 135					
				DB 16840 ATGAACTCTGGAGGAATGATGATATTTCTTCAATTTAAGTAGATGATGGTGGATGGT 16899					
				QY 136 LeuSerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSer 155					
Alignment Scores:				DB 16900 CATCCCAAGTAGACAGTGGTGAATCCCTCGCATATATATCCCATCGCTCACTCACAGC 16959					
				QY 156 Gln----- 156					
				DB 16960 CAGGTTTGTAGAAATTTTGTGATAGCGTGAAGTCGTTAACCCCTCAGGGGTATGTGCAA 17019					
				QY 157 ----- MetSerGlyGluIle 161					
				DB 17020 AGATTGAACATAATTTAGTTTCTTCTCGATTAATCTPATAGATCTCAGGTGAAT 17079					
Pred. No.:				QY 162 ProGlyAlaSerProAspHisMetMetSerProThrGlyAsnIleSerArgAla 181					
				DB 17080 CCTGAGCGCTCCCTGAT---CATATGATGCTCCCGTTGGGAACATTGGCAGACGTGGG 17136					
				QY 182 ---ProPheProTyrValAsnHisSer----- 189					
Score:				DB 17137 CATCCATTCTCCATATGTGAACCATTCAGTAAAGTTTATCTATCATCAAGATCTCATGTTT 17196					
Percent Similarity:									

Alignment Scores:
Pred. No.: 4632.00
Score: 60.06%
Percent Similarity: 60.06%


```
QY 190 -----ProAsnProSerArg 194
Db 17197 CGTTGCATCGACTGGGAAATTTAAATTATTTCTTTGTTGCTCAACAGCAAAACCCATCAAGG 17256
QY 195 GluPheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMet 214
Db 17257 GAGTTTCTCGTAGCTTGGCAATGTTGCATGGAAAGAGAGAGTAGATGCTGGAAATG 17316
QY 215 LysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGly 234
Db 17317 AAG---GATAAAGGTGCAATTTCCATGGCTAAATGGTACTAGCATTTGCTCTCTCAGAAGGT 17373
QY 235 ArgAlaAlaThrAspIleAspAlaSerThrGluTrpAsnMetGluAspAlaLeuLeu--- 253
Db 17374 CGTGGGTTGGTGACATTTGATGATCTACTGACTATAACATGGAGGAGCGCTTACTGTGA 17433
QY 253 ----- 253
Db 17434 GTTCTGATATGCTTTTGTCCACCGTTAGTCATCTATGAATTTTACCAGTAATTAATATTC 17493
QY 253 ----- 253
Db 17494 CTTGTATCTGGTGTACTAACTTGCACTAGTTTGTACTTGTGAAGAACAATAGAAAAATTC 17553
QY 253 ----- 253
Db 17554 TGCATAATTCGAAATATATTTTCTGTTTACAAGTATACCTTGTCTATTAGCAAGGC 17613
QY 253 ----- 253
Db 17614 AATATGCACATTTAATGGATCAGATTTCTAAATGGTTGTACATCTCTGAACCTTTATAATGC 17673
QY 254 -----AsnAspG1 256
Db 17674 GGACTTATCATATAGCTTGCCACTATAAAGTTGTGCTATCTGTATTACAGGAATGATGA 17733
QY 256 uThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTrpAr 276
Db 17734 AACACGCCAACCTCTATCGAGAAAGTGCCAAATTCGTCTATCCAGATTAATCCGTACAG 17793
QY 276 gMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAs 296
Db 17794 GATGGTATTTGTCTCCGATTGATTTGCTATGTATATTCTTGCACCTACCGTATACAAA 17853
QY 296 nProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAl 316
Db 17854 TCTGTGCGTAATGCATACCCGCTATGGTGTCTCTGTCTATATGTAGATTTGGTTGC 17913
QY 316 aLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrIle 336
Db 17914 TCTATCTCGAATCTGGATCAGTTCCCAAGTGGTCCCAATCAACCCGTGAACATTACCT 17973
QY 336 uAspArgLeuAla----- 340
Db 17974 GGATAGGCTGGCTTTAAGGTGATGTACAAATATTTACCTCTGATACACCAATTCCTCGTTA 18033
QY 341 -----LeuArgTyrAspAr 345
Db 18034 GCACGTAGTACTGATTACTGACAATAACAAAATATCATCTTGCATCTCAGGTATGACCG 18093
QY 345 gGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProIle 365
Db 18094 GGAAGGTGAACCATCTCAATGGCTCTCTGTGACATTTTGTGACGTACTGTGATCCCTAT 18153
QY 365 uLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPr 385
Db 18154 GAAAGAACCTCTCTGTGTACGTGCCAAATCTGTCTCTCAATCTTGTGTGGATATACC 18213
QY 385 oValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAl 405
Db 18214 TGTGATAAGGTATCTTGTATGTCTGTATGATGAGCTGCATGTTTAAACATTTTGATGC 18273
QY 405 aLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspI1 425
Db 18274 GCITGTGACACATTTCAGAAATTTGCTAGAAATGGGTACCCGTTCTGTAAAAAGTACACAT 18333
QY 425 eGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysVa 445
Db 18334 AGAACCCAGAGCTCCAGAGTGGTACTTTGCACAGAAAATTTGATTTACTTTGAAGACAAAGT 18393
QY 445 lGlnProSerPheValLysAspArgAlaMetLys----- 457
Db 18394 TCAGGCTTCTTTTGTAAAGATCGTCGTGCCATGAAGGTAAATTTGGTTTATATGTGTGT 18453
QY 457 ----- 457
Db 18454 ACTTTTGTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 18513
QY 458 -----ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuL 476
Db 18514 TCCAGAGGGAATATGAAGAAATTTAAAGTTCTGTTAATGCTCTCGTCGCCAGGCACAGA 18573
QY 476 ysValPr-oGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrA 496
Db 18574 AAGTTCTCTGAGGAAGATGCAATATGCAAGATGGCACACCTTGGCCTGGTAAACAATACCA 18633
QY 496 rgAspHisProGlyMetIle----- 502
Db 18634 GAGACCATCTCGAATGATTCAGGTAATATCACTTAGATATTGCCATTTGTCGTTGTTGTT 18693
QY 503 -----GlnValPheLeuGlyHis 508
Db 18694 ATTGATCCAGCAATTTTCTTACTCTTGACGCTCTGGTCATGCTCAGGTTTCTTGTGTCAT 18753
QY 509 SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 528
Db 18754 AGTGGAGGCTTGTATACCGAAGCAATGAGCTTCTCTGTTTACTCTATGTCTCTCGTGAG 18813
QY 529 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAla-Leu----- 545
Db 18814 AAACGTCCTGGATTCCACACCATAAAAAGGCTGGTGTATGAATGCATTTGTGAGTAGC 18873
QY 545 ----- 545
Db 18874 AGCTAGAGTGGTGTCTGTTAAGTGAGAACTATCTTGTGTTAAATTCATTACTTGCATGTG 18933
QY 546 -----ValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAs 560
Db 18934 GAATTTGAAACTGCAGGTTCTGTTATCAGCTGTCTCTACTAATGGACAATCTTGTGTGA 18993
QY 560 nLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLe 580
Db 18994 TCTTGACTGTGATCACTACATCAACATAGCAAAAGCTCTCCGAGAGGCTATGTGCTTCT 19053
QY 580 uMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspG1 600
Db 19054 TATGATCCAAACCTTAGGAAGGCGTCTGTTATGTCCAATTCCTCAGAGTTTGACCG 19113
QY 600 yIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp----- 616
Db 19114 TATCGATAGAAATGATGATATGCAACAGGAACACCGTGTGTTTTCGATGTAAGCTGTGT 19173
QY 616 ----- 616
Db 19174 TCCCTTTAAGATCTTGGATGCTTCACTTCACGCTTTAGCAACATCAATACTGTTGAT 19233
QY 616 ----- 616
Db 19234 GGATGACCAATATTTATGTTTCTATGTGATGATTTGAACATTTCTGGAGTTTGTCAAAGT 19293
QY 617 -----IleAsnLeuArgGlyLeuAspGlyIleGlnG1 627
Db 19294 AACTTATCTGCATTTCTGCTGCTGCAGATTAATTTGAGAGTCTCGATGGTCTCCAAG 19353
QY 627 yProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluPr 647
```

Db 19354 ACCAGTTTAATGGAACCGGTTGTGTGTTCAACAGACACGCTCTTATGTTATGAACC 19413
Qy 647 oProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAl 667
Db 19414 CCCCATTAAGCAGAAGAGGCGAGGTTACTTCTCTCGCTTTGTGGGGGACGAAAGAGAC 19473
Qy 667 aSerLysSerLysLysArgSerAspLysLysLysSerAsnLysHisValAspSerSe 687
Db 19474 AAAAAAGTCGAAGAGAGAGACGGAAGAAAAAGTCAACAAACATGTGGACAGTTC 19533
Qy 687 rValProValPheAsnLeuGluAspIleGluGlyValGlu 701
Db 19534 TGTGCCAGTTTAAATCTTGAAGATATAGAGAGAGGATGAAGGTACCATCGAGTCGG 19593
Qy 701 701
Db 19594 TTTTGTTCATTAATACATGTTTTTGTCTCTTAGATTTTTAGTGATATTTGGCCACCACCAT 19653
Qy 701 701
Db 19654 GCATGCACCCCAACAACCTTTTTCGAGTATTTCTGTATATGCTTAGGAAGGGAGACA 19713
Qy 702 702
Db 19714 CTAGGCTTAACCTTTTCTTACACTATATATTTGACCTGGCACATCATTTTTGTGGCAGG 19773
Qy 702 yAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPh 722
Db 19774 TTTCTGGATTTGATGATGAGAAATCACTACTGATGTCTCAAAATGAGCTTAGAGAAAAGATT 19833
Qy 722 eGlyGlnSerAlaPheValAlaSerThrLeuMetGluThrGlyGlyValProGlnSe 742
Db 19834 TGGCAATCTAGTGTTTTGTAGCCTCCACTCTGATGGAATATGTGTGTCTTCAATC 19893
Qy 742 rSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyrgluAs 762
Db 19894 CGCAACTCCAGAACTCTTCTGAAGAGGCCATACATGTTATCAGCTGTGGCTAIGAAGA 19953
Qy 762 pLysSerGluTrpGlyThrGlu- 769
Db 19954 CAAAAGTCACTGGGGAACCTGAGGTAATAAGCTATGAGCATTCATGCTCTGCATTAATATC 20013
Qy 769 769
Db 20014 TACTAGCCTACAGTACAAAATTATACAATGTAAACAGAAAAATCGAGATATCAAGTGTC 20073
Qy 769 769
Db 20074 TTTCTGTACAAAATTAAAGGACGCAATTTGATTTCTGTCTGATTCGCACAGTACTTCT 20133
Qy 769 769
Db 20134 GAGTTTCATGCGAAGAAACTTGACTAAATTTGGCATCTCTCTCAGTTGGCAATTAATATTT 20193
Qy 770 770
Db 20194 GTTTATGATGCTATGTGACTTCGAGATTGGATGGATCTATGGTTCTGTTTACAGAAAGATAT 20253
Qy 781 eLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrcysMetProLysAr 801
Db 20254 TCTCACTGGATTCAAGATGATGCAGCGTGGCTGGCGGTCAATCTACTCATGCTCAAGG 20313
Qy 801 gProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuAr 821
Db 20314 ACCAGCAITTCAGGGGTCTGCCCAATTAATCTTTCCGATCGTCTGAACACGAGTGTTCG 20373
Qy 821 gTrpAlaLeuGlySerValGluIleuPheSerArgHisCysProLeuTrpTyrclyTy 841
Db 20374 GTGGGCTCTGTGTTCTGTGGAAATCTTTTTTAGCAGGCATTCCTCCATATGATGAGGTA 20433
Qy 841 rGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrlleAsnThrThrIleTyProle 861

Db 20434 CGGAGGACGGCTTAAGTTCTCGGAGAGATTTGTTTACATCAACACCACCTATTTATCCACT 20493
Qy 861 uThrSerLeuProLeuLeuValTyrcysIleLeuProAlaIleCysIleuLeuThrGlyTy 881
Db 20494 CACATCAATTCACATCTCTTGTACTGCATATGCCAGTATCTGTCTCTCACTGGGAA 20553
Qy 881 sPheIleuMetProGlu- 886
Db 20554 GTTCATCATACCAGAGGTAATATGTAGAAATTTTTCCTCTTCGACAGAAATCGC 20613
Qy 887 887
Db 20614 AGAAGCATCGCTGACTGACTTATTTTGGACTGCAGATTAGCAACTTTGCGAGTATTT 20673
Qy 894 rpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleuGluMetArgTrpSerG 914
Db 20674 GGTTTATCTCTCTCTTCTGTCAATCTTGTACTGTGTATCTTGTAGATGAGGTGAGTG 20733
Qy 914 lyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerA 934
Db 20734 GTGTTGGCATCGATGAATGTTGGAGGATGAGCAGTTTGGGTTATTGGAGGTTATTCG 20793
Qy 934 laHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 954
Db 20794 CTCATCTATTTGCGGCTCTCCAGGCTCTGCTGAAGGTCTTGTGGAATTTGATACAGCT 20853
Qy 954 heThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTrpMetPheL 974
Db 20854 TCCTGTCACTCAAGGCTCTGATGAAGAGGCGATTTCGTGAGCTCTACATGTTCA 20913
Qy 974 ystTrpThrThrLeuLeuIleProThrThrIleLeuIleIleAsnMetValGlyValV 994
Db 20914 AGTGGCAACCTCTTGATCCACCGACCATTTTGTGATCATCAACCTGGTGGTGTG 20973
Qy 994 alAlaGlyThrSerTyrrAlaIleAsnSerGlyTyrglnSerTrpGlyProLeuPheGlyL 1014
Db 20974 TTGCTGGTATCTCTACGCTATCAACAGTGGCTACAGTCATCGGTCCTCTTTGGGA 21033
Qy 1014 ysLeuPhePheAlaPheTrpValIleValHisLeuTyrrProPheLeuLysGlyLeuMetG 1034
Db 21034 AGCTCTCTTTTGGCTTCTGGGTGATTGTCACCTGTACCCCTTCTTAAAGGTCCTCAGG 21093
Qy 1034 lyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuAlaSerIleP 1054
Db 21094 GTGGCAGAACGTCACACCGACCATTTGTGTGTTGGCCATCTCTTGCATCACT 21153
Qy 1054 heSerLeuLeuTrpValArgValAspPropheThrThrArgLeuAlaGlyProAsnIleG 1074
Db 21154 TCTCTTCTGTGGGTTCGTATCGATCCATTCAACCCCGTGTACAGGCCCTGATACCC 21213
Qy 1074 lnThrCysGlyIleAsnCys 1080
Db 21214 AAAATGTGGTATCAACTGC 21233
RESULT 13
AF150630 3723 bp mRNA linear PLN 04-FEB-2000
LOCUS Gossypium hirsutum cellulose synthase catalytic subunit (celA3)
DEFINITION mRNA, complete cds.
ACCESSION AF150630 AF200453
VERSION AF150630.2 GI:6446576
KEYWORDS Gossypium hirsutum (upland cotton)
SOURCE Gossypium hirsutum
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 3723)
Laosinchai,W., Cui,X. and Brown,R.M. Jr.
A full length cDNA of cotton cellulose synthase has high homology
with the Arabidopsis RSW1 gene and the cotton CelA1 gene (Accession
No. AF200453) (PCR 00-002)
Plant Physiol. 122 (1), 291 (2000)
JOURNAL

Db	1415	GCATGTTAACTTTGAAGCACTGCTGAGACATCAGAGTTGCCGGAATGGGTACCT	1474	2495	ATCAGCTGTGGATACGAAGACAAACAGACTGGGGAAGTGAGATTGGGTGGATCTATGGT	2554
Qy	419	PheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIle	438	776	SerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerVal	795
Db	1475	TTCTGCAAGAAATATAACATTGAACCTCGGGCTCCAGAAATGGTATTGTCACAGAAGATT	1534	2555	TTCTGTACAGAAGATATTCTTACTGGATCAAAATGATGCCGTGGTGGCGGTCAATT	2614
Qy	439	AspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArg	458	796	TyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArg	815
Db	1535	GACTACCTGAAGGATAAGTTCAACATCTTTTGTAAAGATCGTAGAGCTATGAAGAGA	1594	2615	TACTGCATGCCAAACGCCAGCCTTAAAGGTTCTGCTCTATCAATCTTTCAGATCGT	2674
Qy	459	GluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValPro	478	816	LeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluLleLeuPheSerArgHisCys	835
Db	1595	GAATATGAGATTTAAAGTTTCGATTAATGCGCTTGTCCGAAGACACAGAAATTCCT	1654	2675	CTGAACCAAGTCTTCGATGGCTCTTGGTCTCTGGAAATCTCTTTTCAGTAGGCATTGC	2734
Qy	479	GluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHis	498	836	ProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsn	855
Db	1655	GAAGAAGGATGGATTATGCAGATGGTACACCATGGCTGCTTAATACACCAAGAGACCAT	1714	2735	CCATCTGGTATGGTTATAGTGGTAGGCTCAAGTGGCTTGAACGATTTCGCATATGTTAAC	2794
Qy	499	ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGlu	518	856	ThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIle	875
Db	1715	CCAGGAATGATCCAGGTTTCTTAGGCCAGAGTGGGGCCCTTGATGCTGAGGGCAATGAG	1774	2795	ACACCATCTACCCAGTCACCTGCTATTCCTCTCTCATGTACTGCACACTGCCAGCAGTC	2854
Qy	519	LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys	538	876	CysLeuLeuThrGlyLysPheIleMetProGluLysSerAsnLeuAlaSerIleTrpPhe	895
Db	1775	CTGCCAAGTTAGTTTATGTTCTCGTGAAGCGTCCAGGCTCCCAACACACACAAAAA	1834	2855	TGCTGCTCTACTAAACAAGTTCAATTATCCACAGATTAGTAACCTTGCAGTATATGGTTT	2914
Qy	539	AlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMet	558	896	IleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyVal	915
Db	1835	GCTGGTGTATGAATGCATCTGTTTCGAGTGTCCAGCAGTCCCTTACTAATGGAGCTTCTTG	1894	2915	ATATCCCTCTTCTCTCATCTTTTGCACCTGGTATTTTAAAAATGAAGTGAATGGTGTG	2974
Qy	559	LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCys	578	916	GlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHis	935
Db	1895	TTAAATCTTGATGTGATCACTACATAAATATAGCAAGGCTTTAAGGGAAGCTATGTGT	1954	2975	GGAATTCACCAATCGTGGAGAAATGAACAGTTTGGGTATCGTGGTGTCTGGCAGAT	3034
Qy	579	PheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPhe	598	936	LeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThr	955
Db	1955	TTTCTGATGGATCCAACTTGGAAAGCAAGTTGTTACGTTCTCCTCAAGGTTT	2014	3035	CTTTTGTGTTTCCAAAGGCTTCTCAAAGTCTCTGCTGGCATTGACACTTAACCTCAG	3094
Qy	599	AspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPheAspIleAsn	618	956	ValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrp	975
Db	2015	GATGGTATTGATCGAATGATCGATATGCCAACAGAAATACTGTTTCTTTGATATAAAC	2074	3095	GTCACTTCAAAAGCATCAGACGAAGATGGGATTTTGTGAACCTTTACATGTTCAAAATGG	3154
Qy	619	LeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsn	638	976	ThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyValValAla	995
Db	2075	TTGAGGGTCTGGATGGCATTCAGGGCCGGTTTATGTCGGTACGGGATGTGTTTCAAT	2134	3155	ACACCCCTTCTCATCCACCAACCACTCTCTCATATAAACTTGGTGGGTGTTGTCA	3214
Qy	639	ArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----LysProGlyPhe	656	996	GlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeu	1015
Db	2135	AGAACAGCATTTGATGGATATGAACCCCTTAAGCCCAAGCATAGAAAAACAGGGATA	2194	3215	GGAATATCGTATGTATCAATAGTGGCTACCAATCATGGGGTCCCCTCTTTGGTAAGCTA	3274
Qy	657	LeuAlaSerLeuCysGlyGly---LysLysLysAlaSerLysSerLysLysArgSerSer	675	1016	PhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArg	1035
Db	2195	CTATCTTCAATTGTCGGAGGTTCTGTAAGAAGAGTTCAAAATCAAGTAAAAAGGATCA	2254	3275	TTCTTGTGCTTCTGGGTGATCATCATCTACCCCTTCTTGAAGAGTTTGAATGGGAGCG	3334
Qy	676	AspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAsp	695	1036	GlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSer	1055
Db	2255	GACAGAAGAATCTGGCAAGCATGTTGATTTCTACTGTACAGATTAATCAATCTAGAGAT	2314	3335	CAAAACCGTACACCTACCATTTGTTGTATGGTCAATCTGCTGTCCTCCATTTTCTCT	3394
Qy	696	IleGluGluGlyValGluGlyValaGlyPheAspAspGluLysSerValLeuMetSerGln	715	1056	LeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThr	1075
Db	2315	ATAGNAGAGGTGTTGAAGGGCTGGGTTTGATGATGAAGTCAATTACTGATGTACAA	2374	3395	TTGTTGGTGGGCGGATCGATCCCTTCAACACAGAGTTACCGGGCCGGATGTAGAACAG	3454
Qy	716	MetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGlu	735	1076	CysGlyIleAsnCys	1080
Db	2375	ATGAGCTGGAGAAACGGTTTGGCCAGTCTGCTGTTTGTGCTCTACACTTATGGAA	2434	3455	TGTGGAATAAACTGT	3469
Qy	736	TyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisVal	755	RESULT 14		
Db	2435	AATGTGTGTCTCCCTCAGTCTCTACCCCGGAACACTTCTTAAGGAAGCAATTCATGTT	2494	BT002335	3229 bp	mRNA linear
Qy	756	IleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGly	775	LOCUS	BT002335	3229 bp mRNA linear
				DEFINITION	Arabidopsis thaliana clone c104938 putative cellulose synthase	linear
				ACCESSION	BT002335	putative cellulose synthase
				VERSION	BT002335.1	mRNA, complete cds.

FLI CDNA.	
Arabidopsis thaliana (thale cress)	
Arabidopsis thaliana	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
1 (bases 1 to 3229)	
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.	
Arabidopsis Open Reading Frame (ORF) Clones	
Unpublished	
2 (bases 1 to 3229)	
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.	
Direct Submission	
Submitted (15-DEC-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	
Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.	
Location/Qualifiers	
1..3229	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/chromosomes="5"
	/clone="C104938"
	/ecotype="Columbia"
	/note="This clone is in puni 51."
1..3229	/gene="At5g05170"
1..3198	/gene="At5g05170"
	/codon_start=1
	/evidence=experimental
	/product="putative cellulose synthase catalytic subunit"
	/protein_id="AA086168.1"
	/db_xref="GI:26983832"
	/translation="MESEGTAGKPMKNIVPQTQICSDNVGKTVGDRFVACDICSF
	PYRCPEYERKDGQSPCKTRVRLKSGSAIPGDKDDGLADGTVFERNVPOKEK
	ISERMLGHLTRGKGEGERPOYDKVEVSHNLPRLTSQDTSGEFSAASPERLSYST
	IAGGRLPYSSDVNGSPNNRIVDPVGLGNVANKERDGMKQKRNTEGVSTQAAASER
	GGVDIDASTDILLADALLDEARQPLSRKVSIPSSRIINPYRMVIMLRVILCLFLHYR
	ITNPVNAFALWSVLCIETWFAISWILDQEPFVFNRETYLDRLALRYDREGPSQ
	LAADVIFVSTVDLPKPLVNTANTVLSILADVDPVKVSCYVSDGAAMLSPELSAET
	SEFARKWVPFCCKYSIEIPRAPEWYFAAKIDYLDKVQTSFVKDRAMKREYEEFKIRI
	NALVSKALKCEEGVMODGTWPNGNTRDHPGMLQVFLGONGGLDAGNELPLRVYV
	SREKEPGQHHKAGAMNALVRVSAVLNPGFPLINDCHYINNKKALREACMFLMDP
	NLGQVCIVQFPQREDDGIDKNRYANRNTVPFDINLRGLDGIOGPVYVGTGCVFNRTA
	LYGVEPPIKVKHKLPSLSLKCGGRKNSKAKESDKKSGRHTDSTVPVFNLDIE
	EGVEGADFDKLLMSQSLERFGQSAFVASTLMENGGVPPSATPENLILKEAIVH
	ISCGVEDKSDMEIGTWYGSVTEDILTGFKMHARGWSIYCMPLKLPAPKSGANLNS
	DLNOLVRLWALGSVILFSRHCPFIWYNGNKLFLERAYVNTVITVPIISPLMYCT
	LPAYCLFNQPIIPQISNIASINFLSFLSIPATGILEMRNSGVIDSWRNBEQFVVI
	GVSVHLFAVQGLDKLVLAGIDNTFTVTSKADEGDFAEFLYFLKWTLLIPPTLLI
	VNLGVAGVSAVINSYGSGMPLFGKLFAPFVIVHLVYFLKGLMGRQNRTPITVVV
	WSVLLASTFSLIMVRIIDPFTSRVTGPDILECGING"
3199..3229	/gene="At5g05170"
3' UTR	
ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	4530.00
Percent Similarity:	87.82%
Best Local Similarity:	78.63%
Query Match:	78.40%
DB:	8
Length:	3229
Matches:	839
Conservative:	98
Mismatches:	104
Indels:	26
Gaps:	8

US-09-900-237A-30 (1-1080) x BT002335 (1-3229)	
QY	19 CysGlnIleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCys 38
DB	58 TGGCAGATCTGTAGTGACAAATGTTGGCAAGACTGTTGATGGAGATCGTTTGTGGCTTGT 117
QY	39 AspValCysArgPheProValCysArgProCysTyrGluHisClnuArglysGluGlyThr 58
DB	118 GATATTGTTCATFTCCAGTGTTCGCGCTTGTACGAGTATGAGAGGAAGATGGGAAT 177
QY	59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArg 78
DB	178 CAATCTGTCTCAGTCGAAACACAGATACAGAGGCTCAAGGTAGTCTGCTATTCCT 237
QY	79 GlyGluGlyAspAspThrAspAlaAspGlySer---AspPheAsnTyrProAla 97
DB	238 GGTGATAAAGACGAGGATGGCTTAGCTGATGAAGGTACTGTTGAGTTCAACTACCCCT--- 294
QY	98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
DB	295 -----CAGAAGGAGAAAATTCAGAGCGGATGCTTGGTTGGCATCTTACT 339
QY	118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
DB	340 CGTGGGAAGGAGAGGAATGGGGAACCCAGTATGATAA----- 381
QY	138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
DB	382 -----GAGGTCTCTCACATCATCTTCTCTCGTCTCAGCAGACAGACAAGAT 426
QY	157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
DB	427 ACTTCAGGAGAGTTTCTGCTGCCTCACCTGAACGGCTCTCTGTATCTTCTACTATCGCT 486
QY	177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
DB	487 GGGGGAAGCGCTTCCCTATTTCATCAGATGTCATCAATCACCATAAGAGGATTGTG 546
QY	196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLys 215
DB	547 GATCCTGTTGGAGTCTGGGAATAGCTTGGAAAGAGAGAGTTCGTGGCTGGAAAAATGAAG 606
QY	216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
DB	607 CAAGAGAAGAACTACTGGTCTCTGTC-----AGCACCGACGCTCTCTCGAA---AGA 654
QY	236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
DB	655 GGTGGAGTAGATATTGATGCCAGCACAGATATCTTAGCAGATGAGGCTCTGCTGAATGAC 714
QY	256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
DB	715 GAAGCGAGCGACGCTCTGTCAAGGAAAGTTTCAATTCCTTCATCAGGATCAATCCTTAC 774
QY	276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
DB	775 AGAATGGTTATTATCTGCGGGCTTTGTTATCTTCTCTTCTTCATCCGATTAACCA 834
QY	296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe 315
DB	835 AACCCAGTGCCAAATAGCCTTCTCTATGGCTGGTCTCTGTCATATGTGAGATCTGGTTT 894
QY	316 AlaLeuSerTrpIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyr 335
DB	895 GCCTTATCTCTGGATTGATCAGTTTCCCAGTGGTTTCTTCGTGAACCGTGAACCCCTAC 954
QY	336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal 355
DB	955 CTCGACAGCGCTTGTCTTAAGATATGATCGTGAAGGTGAGCCATCAGTACGTCTGCTGT 1014
QY	356 AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375

Db 1015 GACATTTTGTGGAGTACTGTTTGACCCCTTGAAGGAGCCACCCCTTGTGACAGCCACACACA 1074
Qy 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395
Db 1075 GTGCTCTCTATTCTGGCTGTGTGACTACCCAGTTGACAAAGGTGCTCTGTATGTTTCTGAT 1134
Qy 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
Db 1135 GATGGTCTCTGTATGTTATCATTTGAATCACTTCAGAAACATCAGAGTTTGCTCGTAAA 1194
Qy 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435
Db 1195 TGGGTACCATTTTGCAGAAATATAGCATAGAGCCTCGTGACACAGAAATGGTACTTTGCT 1254
Qy 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
Db 1255 GCGAAAAATAGATTACTTTGAAGGATAAAGTTTCAGACATCATTTGTCAAAGATCGTAGAGCT 1314
Qy 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
Db 1315 ATGAAGAGGGATATAGGAATTTAAATCCGAATCAATGCACATTGTTTCCAAAGCCCTA 1374
Qy 476 LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr 495
Db 1375 AAATGTCCTGAAGAAGGTGGTATTATCAAGATGGCACACCGTGGCCTGGAAATAATACA 1434
Qy 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
Db 1435 AGGACCATCCAGGAATGATCCAGTCTTCTTAGGGCAAAATGGTGGAGTTGATCGAGAG 1494
Qy 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
Db 1495 GGCAATGAGCTCCCGGTTTGGTATATGTTTCTCGAANAAGCCACAGGATTCAGACAC 1554
Qy 536 HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
Db 1555 CACAAAAGGCTGGTGTATGAATGCATCGTGTAGAGTTTCAGCAGTTCTTTACCAATGGA 1614
Qy 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
Db 1615 CCTTTCATCTTGAATCTTGTATGTGATCATTACATAATAACAGCAAGCCCTTAAGAGAA 1674
Qy 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
Db 1675 GCAATGTCTTCTGATGGACCCAAACCTCGGAAGCAAGTTTGTATGTTTCAGTTCCCA 1734
Qy 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
Db 1735 CAAAGATTTGATGTTATCGATAAGAACGATAGATATGCTAATCGTAATACCGTGTCTT 1794
Qy 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
Db 1795 GATATTAACTTGAGAGGTTTATAGTGGGATTCAGGACCTGTATATGTGCGAACTGGATGT 1854
Qy 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653
Db 1855 GTTTTCAACAGACAGCATATATACGGTTATGACCTCCAAATAAAAGTAAAACACAGAG 1914
Qy 654 ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArg 673
Db 1915 CCAAGTCTTTTATCTAAGCTCTGTGGTGCATCAAGAAAGAGAAATTCCAAAGCTTAAGAA 1974
Qy 674 SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
Db 1975 GAGTCGGCAAAAGAAATACGGCAGGCATATCTGACTCAACTGTTTCTCTGTATTCAACCTC 2034
Qy 694 GluAspIleGluGluGlyValGluGlyValaGlyPheAspAspGluLysSerValLeuMet 713
Db 2035 GATGACATAGAAAGGGAGTTGAAGGTGCTGGTTTGTATGATGATAAGAGCGGCTCTTAATG 2094
Qy 714 SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu 733
Db 2095 TCGCAAAATGAGCCTGGAGAGCGGATTTGGACAGTCTGCTGTTTGTGTTGCTTCTACCCCTA 2154

Qy 734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle 753
Db 2155 ATGAAATAGTGGTGTCTCTCTCAGCAACTCCAGAAACCTTCTCAAGAGGCTATC 2214
Qy 754 HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle 773
Db 2215 CATGTCAATTAGTTGTGTTATGAGATAAGTCAGATTGGGGAATGGAGATTGGATGATC 2274
Qy 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
Db 2275 TATGGTTCTGTGACAGAGATATTCTGACTGGGTTCAAAATGCAATGCCGCGTGGATGCGA 2334
Qy 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
Db 2335 TCCATTTACTGCAAGCCCTAAGCTTCCAGCTTTCAGGGTCTCGCTCTCTATCAATCTTCA 2394
Qy 814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833
Db 2395 GATCGTCTGAACCAAGTGTGAGGTGGGCTTTAGGTTTCAGTTGAGATTCTCTTCAGTCGG 2454
Qy 834 HisCysProLeuThrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
Db 2455 CATTTGCTCTATATGTTATGTTTACAAATGGGAGGCTAAATTTCTTGAGAGGTTTGCCTAT 2514
Qy 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro 873
Db 2515 GTGAACACCAACATCTACCTTATCACCTCTCATCTCTCTTCATGATTGTACATTGCCA 2574
Qy 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
Db 2575 GCGGTTTGTCTCTTCCACCAACCAAGTATTATTCTCAGATTAGTAACATTGCAAGTATA 2634
Qy 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913
Db 2635 TGGTTTCTGCTCTCTCTCTCTCCATTTTCGCCACGGGTATCTAGAAATAGAGTGGAGT 2694
Qy 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
Db 2695 GCGTAGGCATAGACGAATGTGGAGAAACAGCAGGTTTTTGGGTCATTGGTGGAGTATCC 2754
Qy 934 AlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsn 953
Db 2755 GCTCATTTATTCGTGTGTTCATAGGTATCTCTCAAGTCTCTGCGGTATTGACACAAAC 2814
Qy 954 PheThrValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPhe 973
Db 2815 TTCACAGTTACCTCAAAAGCTTCAGATGACAGCAGGACTTTGTCTGAGCTCTACTTGTTC 2874
Qy 974 LysTrpThrThrLeuLeuLeuProThrThrIleLeuIleIleAsnMetValGlyVal 993
Db 2875 AAATGGCAACACTCTGATTCCGCCAACGACGCTCTCATGTAAACTTAGTGGAGATT 2934
Qy 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013
Db 2935 GTTCAGAGGTCCTCTATGCTATCAACAGTGCATCAACCAATCATGGGAGCCACTCTTTGTT 2994
Qy 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
Db 2995 AAGTTGTTCTTTGCTTGTGGTGAITGTTTCACTTGTAACCTTCTCAAGGGTTTGTATG 3054
Qy 1034 GlyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIle 1053
Db 3055 GGTGACAGAAACCGGACTCTTACCATTGTTGGTCTGGTCTGTTCTCTTGGCTTCTATC 3114
Qy 1054 PheSerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
Db 3115 TTCTCGTTGTGTGGTTAGGATTGATCCCTTCTACTAGCCGAGTCACTGGCCCGGACATT 3174
Qy 1074 GlnThrCysGlyIleAsnCys 1080
Db 3175 CTGGAATGTGGAATCAACTGT 3195

RESULT 15
AF027174 3682 bp mRNA linear PLN 07-FEB-1998
LOCUS
DEFINITION Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B)
MRNA, complete cds.
ACCESSION AF027174
VERSION AF027174.1 GI:2827142
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 3682)
Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W.,
Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J.,
Redmond, J. and Williamson, R.E.

Molecular analysis of cellulose biosynthesis in Arabidopsis
Science 279 (5351), 717-720 (1998)
98111412
9445479

2 (bases 1 to 3682)
Arioli, T.
Direct Submission
Submitted (29-SEP-1997) Plant Science Centre, Australian National
University, Acton, Canberra, ACT 200, Australia
Location/Qualifiers

1. 3682
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
1. 3682
/gene="Ath-B"
256..3453
/gene="Ath-B"
/note="RSW1-like"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAC39336.1"
/db_xref="GI:2827143"
/translation="MESGETAGKPMKNIIVPTQICISDNVGTVDGRFVACDICSF
PVCRCYERKDGNSQPCQKRYRLKGSIPAIFGDKDEGLADEGTFVFNYPQKEK
ISERMLGHLTRKGEEGEPQYDKREVSHNLPLRTSRQDTSGEFSAPRLSVSST
IAGKRLPYSSDVNOSPNNRI VDPVGLGNVAMKERVGDGKMKOEKNTGPTQAAASR
GGVDIDAGTDILADALLDEARQPLSRKVSIPSSR INPYRWIMLRVLVILFLHYR
ITNPVNAFALWLSVI CEIWFAL SWILDQPKFPVNRREYDLRLALRYDREGPSQ
LAAVDIFVSTDPLEKEPLVTVLSILAVDIPVDRKVSIVFDDGAWLSPESLAET
SEFARKWVFFCKKYSIEPRAPWYFAAKIDYLKQVTSFVKDRAMKREYEEFKIRI
NALSKALKCEEGWVMDQGTWPNGNTGDHPGMIQVFLGQNGLDAREGNELPLRVY
SREKRFQHHKAGAMNALVRSAVLNTGPFILNLDGDHYINNSKALREAMCFLMDP
NLGKQVCYVQPPQFDGIDKNDYANRNTVPFDINLRGLDGIQGVVYVGTGVENRTA
LYGVEPPIKVXHKHESLKLCCGSRKKNKAKESKKKSGRHHDSIVPVNLDIE
EGVEGAGDEDEKALLMSQMSLEKRFQGSVAVFASVTLMENGVPVPSPATPENLLKEAIV
ISCGVEDSKDMEIGWIGVSTEDILTFQKHARGRSIYCMKPLPAFGSAPINLS
DRLNQLRWALGSILFESRHCPWIWYNGNRLKFLERFAYNTVITPTISPLMYCT
LLAVCLFTNQFIPOISNIAISWFLSLFSLFATGILBMRWSGVGIDSWRNEQFWVI
GGVSAHLFAVQGIKLVAGIDTNFTVTSKASDEGDFAEYLFPKWTLLIPPTLLI
VNLGVVAGVSYAINSGQSMCPFGKLPFAFWIVLHYLPFLKGLMGRQNRTPIVVV
WSVLLASTFSLNWRIDFTSRVTGDFDILEGGINC"

Alignment Scores:
Pred. No.: 0 Length: 3682
Score: 4507.00 Matches: 836
Percent Similarity: 87.54% Conservative: 96
Best local Similarity: 78.35% Mismatches: 107
Query Match: 78.00% Indels: 26
DB: 8 Gaps: 8
US-09-900-237a-30 (1-1080) x AF027174 (1-3682)

ORIGIN

313 TGCCAGATCTGTAGTGACAAATGTTGGCAAGACTGTTGTATGGAGATCGTTTGTGGCTGT 372
39 AspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThr 58
373 GATATTGTTTCATTCCTCCAGTGTTCGCGCCTTGTACGAGTATGAGAGGAAGATGGAAT 432
59 GlnAlaCysLeuGlnCysLysThrLysTyrLysAsqHisArgLysSerProAlaIleArg 78
433 CAATCTTGCTCAGTGCAGAAACCCAGATACAAGAGCTCAAAGGTAGTCCTGTATTCCT 492
79 GlyGluGluGlyAspThrAspAlaAspAspGlySer---AspPheAsnTyrProAla 97
493 GGTGATAAAGACAGCAGGTGGCTTAGCTGATGAAGTACTGTTCAGTTCACATACCCT--- 549
98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
550 -----CAGAAGGAGAAAATTTTCAGACGGATGCTTGTGTTGGCATCTTACT 594
118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
595 CGTGGAGGAGAGAAATGGGGAAACCCCAAGTATGATAA----- 636
138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
637 -----GAGGTCTCTCAATCATCTCTCGTCTCAGCAGACAGACAGAT 681
157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
682 ACTTCAGGAGAGTTCCTGCTGCCTCACCTGAACGCCTCTCTGATCTTCTACTACGCT 741
177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
742 GGGGAAAGCGCTTCCTTATTCATCAGATGTCATCATCAATCAACCAATAGAAGGATTGTG 801
196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLys 215
802 GATCTGTTGGACTCGGGAATGTAGCTTGGAAAGAGAGAGTGTGCTGGTGGAAAATGAAG 861
216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
862 CAAGAGAAGAACTACTGTCCTGTC-----AGCAGCAGCGCTGCTCTGAA---AGA 909
236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
910 GTGAGTAGATATTGATGCCAGCACAGATATCTCAGCAGATGAGCTCTGCTGATGAC 969
256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
970 GAAGCGAGCAGCCTCTGTCGAAGGAAAGTTTCAATTCCTTCATCAGGATCAATCCTTAC 1029
276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
1030 AGAATGGTATTATATGCTCGCGCTTGTATCTCTCTCTCTTCATTCACCGTATAACA 1089
296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysLulileTrpPhe 315
1090 AACCCAGTGCCAATGCCTTGTCTATGGCTGGTCTCTGTGATATGAGATCTGGTTT 1149
316 AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr 335
1150 GCCTTATCTCGATTTTGGATCAGTTTCCCAAGTGGTTTCTCTGTGAACCGTGAAACCTAC 1209
336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaIleVal 355
1210 CTCGACAGCTTGCTTTAAGATATGATCGTGAAGGTGAGCCATCAGATGACTGCTGTT 1269
356 AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375
1270 GACATTTTCGTGAGTACTGTTGACCCCTTGAGGAGGCCACCCCTTGTACAGCCACACA 1329
376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395

19 CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys 38

Db 1330 GTGCTCTCTATTCTGGCTGTGACTACCCAGTTGACAAGGTGTCCTGTTATGTTTTTGAT 1389
 QY 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
 Db 1390 GATGGTGCTGCTATGTTATCATTTGAATCACTTGCAGAAACATCAGAGTTTGTCTGTCTAAA 1449
 QY 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435
 Db 1450 TGGGTACCACTTTTGCAGAAATATAGCATAGAGCCCTGTCACAGAAATGGTACTTTGCT 1509
 QY 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
 Db 1510 GCGAAATAGATTACTTGAAGGATAAAGTTCCACATCATTTCTCAAAGATCGTAGCT 1569
 QY 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
 Db 1570 ATGAGAGGGAATATGAGGAATTTAAATCCGAATCAATGCACTTGTCTTCAAAGCCCTA 1629
 QY 476 LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr 495
 Db 1630 AATGTCTCGAAGAGGGTGGTTATGCAAGATGGCACACCGTGGCTCGAATAATATACA 1689
 QY 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
 Db 1690 GGGGACCATCCAGGAATGATCCAGTCTCTTAGGGCAAAATGGTGACTTGTATGCAGAG 1749
 QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
 Db 1750 GGCAAATGAGCTCCGGCTTGGTATATGTTCTTCAGAAAAGGACCCAGGATTCACGAC 1809
 QY 536 HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
 Db 1810 CACAAAAGGCTGGTGTATGAATGCACTGGTGAGAGTTTCAGCAGATTCITACCAATGGA 1869
 QY 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
 Db 1870 CCTTTCACTCTTGATCTTGATTGTGATCAATTACATAATAACAGCAAGCCTTAAGAGAA 1929
 QY 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
 Db 1930 GCAATGTCTTCTGTATGGACCCAAACCTCGGAAGCAAGTTTGTATGTTCACTTCCCA 1989
 QY 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
 Db 1990 CAAAGATTGTATGGTATCGATAAGAACGATAGATATGTAATCGTAATACCGTGTCTTT 2049
 QY 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
 Db 2050 GATATTAACTTGAGAGTTTAGATGGGATTCAGGACCTGTATATGTCGGAACCTGGATGT 2109
 QY 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653
 Db 2110 GTTTTCAACAGAACAGCATTTATACGGTTATGAACCTCCAATAAAGTAAACACACAGAAG 2169
 QY 654 ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArg 673
 Db 2170 CCAAGTCTTTATCTAAGCTCTGTGGTGATCAAGAAAGAAATTCCAAGCTTAAGAA 2229
 QY 674 SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
 Db 2230 GAGTCGGACAAAGAAATACGGCAGGCATACCTGACTCAACTGTTCTCTGATTCAACCTC 2289
 QY 694 GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet 713
 Db 2290 GATGACATAGAAGAGGAGTTGAAGGTGCTGTTTGTATGATGAAAAAGCGCTCTTAATG 2349
 QY 714 SerGlnMetSerLeuLysLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu 733
 Db 2350 TCGCAAAATGAGCCGTGGAGAGCGATTTGGAGAGTCTGCTGTTTGTGCTTCTACCCCTA 2409
 QY 734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle 753
 Db 2410 ATGGAATATGTTGTTCTCTCTCTCAGCAACTCCAGAAAACCTTCTCAAAGAGGCTATC 2469

QY 754 HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle 773
 Db 2470 CATGTCAATTAGTTGTGGTTATGAGGATAAGTCAGATTGGGAATGGAGATTGGATGCATC 2529
 QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
 Db 2530 TATGGTTCTGTGACAGAGATATTCAGCTGGGTTCAAAATGTCATGCCGTGGATGGCGA 2589
 QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
 Db 2590 TCCATTACTGTCATGCTTAAGCTTCCAGCTTTCAGGGTTCTGCTCTCATCAATCTTTCA 2649
 QY 814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833
 Db 2650 GATCGTCTGACCAAGTGTGAGTGGCTTTAGGTTTCAGTTGAGATTCTCTTCAGTCGG 2709
 QY 834 HisCysProLeuThrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
 Db 2710 CATTTGCTCTATATGGTTATGGTTACAATGGGAGCTAAATTTCTTGAGAGGTTTGGCTAT 2769
 QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro 873
 Db 2770 GTGAACACCAACCATCATCCCTATCCTCCTCTCTCATGATTGTGACATTGCTA 2829
 QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
 Db 2830 GCCGTGTGTCTCTTACCAACCATGTTTATTTCTCCTCAGATTAGTAACATTGCAAGTATA 2889
 QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrPhe 913
 Db 2890 TGGTTTCTGTCTCTCTTCTCTCCATTTTCGCCACGGGTATACTAGAAATGAGGTGGAGT 2949
 QY 914 GlyValGlyIleAspGluTyrPheArgAsnGluGlnPheThrValIleGlyIleSer 933
 Db 2950 GGGGTAGGCATAGCAATGTTGGAGAAACGACAGCTTTTGGGTCAITGGTGAGTATCC 3009
 QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsn 953
 Db 3010 GCTCATTTATTCCTGTGTGTTTCAAGGTATCTCTCAAGTCTCTGCCGTATTGACAAAC 3069
 QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
 Db 3070 TTCACAGTTACCTCAAAAGCTTCAGATGAAGACGAGACTTTGCTGAGCTCTACTTCTTC 3129
 QY 974 LysThrThrThrLeuLeuIleProProThrThrThrIleLeuIleIleAsnMetValGlyVal 993
 Db 3130 AAATGGACAAACACTTCTGTATTCGCCCAACGACGCTCTCATTTGTAACCTTAGTGGGAGTT 3189
 QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013
 Db 3190 GTTGACAGAGTCTCTATGTATCAACAGTGGATACCAATCATGGGACCACTCTCTTGGT 3249
 QY 1014 LysLeuPhePheAlaPheThrValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
 Db 3250 AAGTTGTTCTTTGGCTTCTGGGTGATTGTTCACTGTACCTTTCTCAAGGGTTTGTATG 3309
 QY 1034 GlyArgGlnAsnArgThrProThrIleValIleValThrAlaValLeuLeuAlaSerIle 1053
 Db 3310 GGTGACAGAAACCGGACTCTCTACCATTTGTTGTGGTCTGGTCTGTCTTCTTGGCTTCTATC 3369
 QY 1054 PheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
 Db 3370 TTTCTGTTGTTGTTGGTTAGGATTGATCCCTTCACTAGCCGAGTCACCTGCCCCGACATT 3429
 QY 1074 GlnThrCysGlyIleAsnCys 1080
 Db 3430 CTGGAATGTGAATCAACTGT 3450

Search completed: August 22, 2004, 20:43:20
 Job time : 10363 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2004, 22:21:58 ; Search time 92 Seconds
(without alignments)

3316.862 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MGDADALSKRGHAGDVQ.....VDFFTTLACPNITQCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: geneseqp1980s.*

2: geneseqp1980s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5423.5	93.9	1079	5 AAU10496	AAU10496 Corn cell
2	5415.5	93.7	1077	3 AAY84108	AAY84108 Amino aci
3	5415.5	93.7	1077	3 AAY84114	AAY84114 Amino aci
4	5255.5	91.0	1043	3 AAY84119	AAY84119 Amino aci
5	5143	89.0	1076	3 AAY84121	AAY84121 DNA encod
6	4530	78.4	1054	3 AAG48734	AAG48734 Arabidops
7	4530	78.4	1065	3 AAG48733	AAG48733 Arabidops
8	4530	78.4	1065	3 ABB93356	ABB93356 Herbicida
9	4499	77.9	1065	2 AAU33819	AAU33819 Arabidops
10	4209.5	72.9	959	3 AAG48735	AAG48735 Arabidops
11	3993.5	69.1	1074	3 AAY84110	AAY84110 Amino aci
12	3993.5	69.1	1074	3 AAY84112	AAY84112 Amino aci
13	3993.5	69.1	1074	3 AAY84117	AAY84117 Amino aci
14	3981	68.9	1075	3 AAY84120	AAY84120 Amino aci
15	3981	68.9	1075	3 AAY84115	AAY84115 Amino aci
16	3981	68.9	1075	3 AAY84109	AAY84109 Amino aci
17	3963.5	68.6	1081	2 AAU33817	AAU33817 Arabidops
18	3963.5	68.6	1081	5 ABB93199	ABB93199 Herbicida
19	3953.5	68.4	1081	2 AAU33820	AAU33820 Arabidops
20	3844.5	66.5	1094	3 AAY84118	AAY84118 Amino aci
21	3844.5	66.5	1094	3 AAY84113	AAY84113 Amino aci
22	3844.5	66.5	1165	3 AAY58834	AAY58834 Corn cell
23	3820.5	66.1	1039	3 AAY58837	AAY58837 Soybean c
24	3820	66.1	1058	3 AAY84107	AAY84107 Amino aci
25	3818	66.1	1086	3 AAY58835	AAY58835 Corn cell

26	3818	66.1	1086	3 AAY84111	AAY84111 Amino aci
27	3818	66.1	1086	3 AAY84116	AAY84116 Amino aci
28	3818	66.1	1148	3 AAY58832	AAY58832 Corn cell
29	3801	65.8	1026	5 ABB93522	ABB93522 Herbicida
30	3778.5	65.4	1065	5 ABB91880	ABB91880 Herbicida
31	3777.5	65.4	1039	2 AAU73309	AAU73309 Cellulose
32	3746	64.8	1043	5 ABB93701	ABB93701 Herbicida
33	3666.5	63.5	1084	5 ABB93304	ABB93304 Herbicida
34	3660.5	63.4	1069	5 ABB93411	ABB93411 Herbicida
35	3636	62.9	1084	5 ABB93949	ABB93949 Herbicida
36	3633.5	62.9	1084	2 AAU33818	AAU33818 Arabidops
37	3579	61.9	1088	5 ABB91831	ABB91831 Herbicida
38	3566.5	61.7	1080	4 AAB37895	AAB37895 Arabidops
39	3535	61.2	821	5 AAU10495	AAU10495 Corn cell
40	3482	60.3	974	2 AAU73308	AAU73308 Cellulose
41	3477	60.2	974	2 AAU60039	AAU60039 Cotton ce
42	3408.5	59.0	955	3 AAB16338	AAB16338 Eucalyptu
43	3302.5	57.2	958	5 ABB92981	ABB92981 Herbicida
44	3234	56.0	793	3 AAY58839	AAY58839 Soybean c
45	3157	54.6	946	4 AAB37893	AAB37893 Populus t

ALIGNMENTS

RESULT 1

AAU10496

ID AAU10496 standard; protein; 1079 AA.

XX AAU10496;

XX AC

XX 14-FEB-2002 (first entry)

XX DE Corn cellulose synthase Cqrae19/cesa-9.

XX KW Corn; cellulose synthase; Cdpgs45; cesa-3; Cqrae19; cesa-9;

XX KW stalk quality; improved stand; silage; pericarp; kernel hardening;

XX KW handling ability; transgenic plant; immunogen.

XX OS Zea mays.

XX PN WO200179516-A2.

XX PD 25-OCT-2001.

XX PF 12-APR-2001; 2001WO-US011951.

XX PR 14-APR-2000; 2000US-00550483.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Dhugga KS, Helentjaris TG;

XX DR WPI; 2002-041338/05.

XX N-PSDB; AAS16458.

XX PT New cellulose synthase polypeptides and polynucleotides, useful in

XX PT improving stalk quality or silage, and in increasing concentration of

XX PT cellulose in the pericarp, hardening the kernel for improved handling

XX PT ability.

XX PS Claim 3; Page 85-87; 88pp; English.

XX CC The invention relates to isolated nucleic acids encoding two cellulose

XX CC synthase proteins from corn, Cdpgs45 (cesa-3) and Cqrae19 (cesa-9). Also

XX CC disclosed are a recombinant expression cassette comprising the

XX CC polynucleotide (operably linked to a promoter) a host cell comprising the

XX CC recombinant expression cassette and a transgenic plant comprising the

XX CC recombinant expression cassette. The nucleic acid is useful in the

XX CC improvement of stalk quality for improved stand or silage, and in the

XX CC increased concentration of cellulose in the pericarp, hardening the

XX CC kernel, and thus improving its handling ability. The nucleic acids may

XX CC also be used as probes or amplification primers in the detection,

CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 1077 AA;

Query Match 93.7%; Score 5415.5; DB 3; Length 1077;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 997; Conservative 46; Mismatches 34; Indels 3; Gaps 2;

QY 1 MDGADALKSGRHGAGDVCQICADGLGTLTDGVDFTACDVCRCPPVCYEHKKEGTQA 60
DB 1 MEGDADGVKSGRRGGQVCQICGQGVGTAGEDVFAACDVCQFPVCYEHKKEGTQA 60
QY 61 CLQCKTKYKRGHSGPAIRGEEDGTDADGSDFNYPASGTEDKOKIADRMRSWRMTGG 120
DB 61 CPQCKTKYKRGHSGPAIRGEEDGTDAD--SDFNYLASGNEQKOKIADRMRSWRMTGG 118
QY 121 SGNVGHEKYSGEIGLKYDSGEIPRGVPSVTSQMSGEIPGASPDHMMSPGNTSRR 180
DB 119 SGDVGRPKYDSGEIGLKYDSGEIPRGVPSVTSQMSGEIPGASPDHMMSPGNTIGKR 178
QY 181 APFFVNVHSPNPSREFSGISGNVAKKRVKQKQKGAIPMTNGTSIAPSEGRGAATDI 240
DB 179 APFFVNVHSPNPSREFSGISGNVAKKRVKQKQKGTIPMTNGTSIAPSEGRGVCDI 238
QY 241 DASTYNNMEDALLNDETQPLSRKVPKPIASSKINPYRMVILRLVLSIFLHYRTNPNRN 300
DB 239 DASTYNNMEDALLNDETQPLSRKVPKPIASSKINPYRMVILRLVLSIFLHYRTNPNRN 298
QY 301 AYPILWLSVCEIWFALSWILDQPKFPINRETYLDELALRYDREGEPSQLAAVDIFVS 360
DB 299 AYPILWLSVCEIWFALSWILDQPKFPINRETYLDELALRYDREGEPSQLAAVDIFVS 358
QY 361 TVDELKPEPIVANTVLSILAVDPVKVCSYVSDGASMLTFDALAETSFARKKWPFFV 420
DB 359 TVDEPKPEPIVANTVLSILAVDPVKVCSYVSDGASMLTFDALAETSFARKKWPFFV 418
QY 421 KKYDIEPRAPPEFYFCQIDYLDKQVPSFVKDRAMKREVEEFKIRINALVSKALKVPEE 480
DB 419 KKYDIEPRAPPEFYFCQIDYLDKQVPSFVKDRAMKREVEEFKIRINALVSKALKVPEE 478
QY 481 GWIMODGTPGNNTRDPHGMIOVFLHSGSLDTEGNELPRLVYVSREKRGFOHKKAG 540
DB 479 GWIMODGTPGNNTRDPHGMIOVFLHSGSLDTEGNELPRLVYVSREKRGFOHKKAG 538
QY 541 ANNALVRVSAVLITNGQYMLNDCDHYINNSKAVREAMCFLMDPNLGPQVCYVQPPQPDG 600
DB 539 ANNALVRVSAVLITNGQYMLNDCDHYINNSKAVREAMCFLMDPNLGRSVCYVQPPQPDG 598
QY 601 IDRNDRVANRNTVFDINLRGLDGIQGVVYVGTGCVFNRTAIYGYEPPKAKPGFTLASL 660
DB 599 IDRNDRVANRNTVFDINLRGLDGIQGVVYVGTGCVFNRTAIYGYEPPKAKPGFTLASL 658
QY 661 CGGKKKASKRSKRSKSKKSKKHVDSSVPVFNLEDEIEGVEGAGFDDKSKVLSQMSLEK 720
DB 659 CGGKKKASKRSKRSKSKKSKKHVDSSVPVFNLEDEIEGVEGAGFDDKSKVLSQMSLEK 717
QY 721 RFGQSAFVASTLMEYGVGPQSSPTESLLKEAHIVISCGYEDKSEWGPTEIGWYGSVTE 780
DB 718 RFGQSAFVASTLMEYGVGPQSSPTESLLKEAHIVISCGYEDKSEWGPTEIGWYGSVTE 777
QY 781 ILTGFKMHARGWRSVYCMKPKPAFKGSAPINLSRLNQVLWALGSVEILFSRHCPMYG 840
DB 778 ILTGFKMHARGWRSVYCMKPKPAFKGSAPINLSRLNQVLWALGSVEILFSRHCPMYG 837
QY 841 YGGRKFLERFAYINTTYPITSLPLLYCYLPAICLLTGKFIEMPEINSLASINFIALFL 900
DB 838 YGGRKFLERFAYINTTYPITSLPLLYCYLPAICLLTGKFIEMPEINSLASINFIALFL 897
QY 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFAVFGLLKVLGADITNFTVTSKA 960

DB 898 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFAVFGLLKVLGADITNFTVTSKA 957
QY 961 NDEGDPAELYMFKWTTLLPPTTILLINMGVVGAGTSYAINSGYQSWGPLFGKLFPAFW 1020
DB 958 SDEGDPAELYMFKWTTLLPPTTILLINMGVVGAGTSYAINSGYQSWGPLFGKLFPAFW 1017
QY 1021 VIVHLYPFLKGLMGRQNRTPITIVWAVLLASIPSLWVRVDPPTTTLRAGNIQTCCGINC 1080
DB 1018 VIVHLYPFLKGLMGRQNRTPITIVWAVLLASIPSLWVRVDPPTTTLRAGNIQTCCGINC 1077

RESULT 3

AAY84114

ID AAY84114 standard; protein; 1077 AA.

XX AC AAY84114;

XX AC AAY84114;

DT 03-JUL-2000 (first entry)

XX Amino acid sequence of a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker.

XX Zea mays.

XX Key Location/Qualifiers

FT Misc-difference 494

FT /note= "encoded by MGG"

XX W0200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US018760.

XX 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI: 2000-224343/19.

XX N-PSDB; RAZ99512.

XX New genes which encode maize cellulose synthase polypeptides in plants

XX useful for modulating the expression of cellulose synthase in plants and

XX to produce transgenic plants expressing the novel protein.

XX Claim 15; Page 142-144; 119pp; English.

XX The present sequence represents a maize cellulose synthase polypeptide.

XX The cellulose synthase can be used for the improvement of stalk quality

XX for improved stand or silage. It also provides an increased concentration

XX of cellulose in the pericarp, hardening the kernel and improving its

XX handling ability. The sequences are used to produce transgenic plants and

XX seeds expressing the cellulose synthase. The polynucleotide is used for

XX modulating, preferably increasing, the level of the synthase in a plant

XX cell. The plants are preferably monocots. The polynucleotide is also used

XX as a probe or primer in the detection quantitation or isolation of gene

XX transcripts. The probes are useful in detecting deficiencies in the level

XX of mRNA in screenings for desired transgenic plant, for detecting or

XX mutations in the gene, for monitoring upregulation of expression or

XX changes in enzyme activity in screening assays of compounds, for

XX detection of any number of allelic variants of the gene, or for use as

XX molecular markers in plant breeding programs. The isolated nucleic acids

XX of the present invention can also be used for recombinant expression of

XX their encoded polypeptides or for use as immunogens in the preparation

XX and/or screening of antibodies. The proteins can be employed in assays

XX for enzyme agonists or antagonists of enzyme function or for use of

XX immunogens or antigens to obtain antibodies specifically immunoreactive

XX with a protein

SQ	Sequence 1077 AA;	
	Query Match 93.7%; Score 5415.5; DB 3; Length 1077;	
	Best Local Similarity 92.3%; Pred. No. 0;	
	Matches 997; Conservative 46; Mismatches 34; Indels 3; Gaps 2;	
QY	1 MDGDADALSKSRHGAGDVQICADGLGTLTLDGVDFTACDVCRCPPVCRPCYEHKREGTQA 60	
DB	1 MEGDADGVKSGRRGGQVQICDGVGTTAEGDVAACDVCGFPVCRPCYEVKRDGTQA 60	
QY	61 CLOCTKYKRRHSGSPAIRGEGDDTDADGSDPNYPASGTEDQKQKIADMRMRWNTGG 120	
DB	61 CPQCKTKYKRRHSGSPAIRGEGDDTDAD--SDFNYLASGNEQKQKIADMRMRWNVGG 118	
QY	121 SGNVGHPKYSDEIGLSKYDSGEIPRGVYVPSVTNSQMSGEIPGASPDHMMSPGTGNISR 180	
DB	119 SGVGRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMMSPGTGNIGKR 178	
QY	181 APPFYVNSPNSPREFSGSIGNVAMKERVDMGKMKQDKGAIEMTNGTSLAPSEGRAATDI 240	
DB	179 APPFYVNSPNSPREFSGSIGNVAMKERVDMGKMKQDKGTIEMTNGTSLAPSEGRGVGDI 238	
QY	241 DASTYNNMEDALLNDETROPILSRKVPDIASSKINPRMVIIVLVLVLSIFLHYRLNTPVN 300	
DB	239 DASTYNNMEDALLNDETROPILSRKVPDIASSKINPRMVIIVLVLVLSIFLHYRLNTPVN 298	
QY	301 AYPLMLLSVICBIWFALSWILDQFPKPPINRETYLDRALRYDREGESQLAADVFS 360	
DB	299 AYPLMLLSVICBIWFALSWILDQFPKPPINRETYLDRALRYDREGESQLAADVFS 358	
QY	361 TVDPLKEPPIVNTANTVLSILAVDYPVDKVCYSVSDGASMLTFDALAETSEFARKWPFV 420	
DB	359 TVDPMKEPPLVNTANTVLSILAVDYPVDKVCYSVSDGGAAMLTFDALAETSEFARKWPFV 418	
QY	421 KYDIEPRAPERYFOCKDYLKDKQVPSFKDORRAMKREYEFKIRINALVSKALKVPEE 480	
DB	419 KXNIEPRAPERYFOCKDYLKDKVHPFSFKDORRAMKREYEEFKVRNGLVAKAKVPEE 478	
QY	481 GWIMQDGTWPGNNTDRHPGMIQVFLHSGGGDTEGNELPRLVYVSREKRPQFQHKKAG 540	
DB	479 GWIMQDGTWPGNNTXDRHPGLQVFLHSGGLDTGNEMLPRLVYVSREKRPQFQHKKAG 538	
QY	541 ANNALVRVSAVLTNGQYMLNDCDHYYINNSKAVREAMCFMDPNLGPQVYVQFPQREDG 600	
DB	539 ANNALVRVSAVLTNGQYMLNDCDHYYINNSKALREAMCFMDPNLGRSVYVQFPQREDG 598	
QY	601 IDRDYANRNTVFFDINLRGLDGIQGVYVGTGCVENRTALGYEPPIKAKPKFLASL 660	
DB	599 IDRDYANRNTVFFDINLRGLDGIQGVYVGTGCVENRTALGYEPPIKQKKGFLSSL 658	
QY	661 CGGKKASKSKRSRSDKKSKNKHVDSVPVFNLEDEIEGVEGAGFDDDEKSVLMSQMSLEK 720	
DB	659 CGGRKASKSKK-GSDKKSKKHVDSVPVFNLEDEIEGVEGAGFDDDEKSVLMSQMSLEK 717	
QY	721 RFGQAAVASTLMYGYGVSQSTPESLLKEAHIVISCGYEDKSWGTEIGWYGSVTE 780	
DB	718 RFGQAAVASTLMYGYGVSQSTPESLLKEAHIVISCGYEDKTEWTEIGWYGSVTE 777	
QY	781 ILTGPFMHARGWRSVYCMKPAFKGAPINLSDRNLQVLRWALGSVELSRHCPMYG 840	
DB	778 ILTGPFMHARGWRSVYCMKPAFKGAPINLSDRNLQVLRWALGSVELSRHCPMYG 837	
QY	841 YGGRKFLERFAYINTTIIYPLTSLPLLVYLCILPAICLTGTFIMPEINSLASIWIALFL 900	
DB	838 YGGRKFLERFAYINTTIIYPLTSLPLLVYLCILPAICLTGTFIIEISNFIWISLFI 897	
QY	901 SIFATGILEMRWSGVGIDEWNRNEQFVIGGISAHLPFAVFGQLLKVLAGIDNFTVTSKA 960	
DB	898 SIFATGILEMRWSGVGIDEWNRNEQFVIGGISAHLPFAVFGQLLKVLAGIDNFTVTSKA 957	
QY	961 NDEEDGDAELYNFKWTTLLIPPTTILIIINLVGVVAGTGYAINSGYQSWGPLFGKLPFAFW 1020	
DB	958 SDEEDGDAELYNFKWTTLLIPPTTILIIINLVGVVAGISYAINSGYQSWGPLFGKLPFAFW 1017	

QY	1021 VIVHLYPFLKGLMGRQNRTPETIVVAVLLASIFSLLVWRVDPFTTTLRAGPNITQCGINC 1080	
DB	1018 VIVHLYPFLKGLMGRQNRTPETIVVAVLLASIFSLLVWRVDPFTTTRVTPGTQTCGINC 1077	
RESULT 4		
AY84119		
ID	AY84119 standard; protein; 1043 AA.	
AC	AY84119;	
XX	03-JUL-2000 (first entry)	
XX	Amino acid sequence of a maize cellulose synthase.	
XX	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;	
KW	transgenic plant; plant breeding marker.	
XX	Zea mays.	
OS		
XX		
PH	Key Location/Qualifiers	
FT	Misc-difference 494	
FT	/note= "encoded by MGG"	
XX		
PN	W020009706-A2.	
XX		
PD		
XX		
PF	16-AUG-1999; 99WO-US018760.	
PR	17-AUG-1998; 98US-0096822P.	
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;	
XX		
DR	WPI; 2000-224343/19.	
XX	N-PSDB; AAZ99527.	
PT	New genes which encode maize cellulose synthase polypeptides in plants	
PT	useful for modulating the expression of cellulose synthase in plants and	
XX	to produce transgenic plants expressing the novel protein.	
PS	Claim 15; Page 181-183; 119pp; English.	
XX		
CC	The present sequence represents a maize cellulose synthase polypeptide.	
CC	The cellulose synthase can be used for the improvement of stalk quality	
CC	for improved stand or silage. It also provides an increased concentration	
CC	of cellulose in the pericarp, hardening the kernel and improving its	
CC	handling ability. The sequences are used to produce transgenic plants and	
CC	seeds expressing the cellulose synthase. The polynucleotide is used for	
CC	modulating, preferably increasing, the level of the synthase in a plant	
CC	cell. The plants are preferably monocots. The polynucleotide is also used	
CC	as a probe or primer in the detection quantitation or isolation of gene	
CC	transcripts. The probes are useful in detecting deficiencies in the level	
CC	of mRNA in screenings for desired transgenic plant, for detecting or	
CC	mutations in the gene, for monitoring upregulation of expression or	
CC	changes in enzyme activity in screening assays of compounds, for	
CC	detection of any number of allelic variants of the gene, or for use as	
CC	molecular markers in plant breeding programs. The isolated nucleic acids	
CC	of the present invention can also be used for recombinant expression of	
CC	their encoded polypeptides or for use as immunogens in the preparation	
CC	and/or screening of antibodies. The proteins can be employed in assays	
CC	for enzyme agonists or antagonists of enzyme function or for use of	
CC	immunogens or antigens to obtain antibodies specifically immunoreactive	
CC	with a protein	
XX		
SQ	Sequence 1043 AA;	

Query Match 91.0%; Score 5255.5; DB 3; Length 1043;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 969; Conservative 42; Mismatches 32; Indels 3; Gaps 2;

QY 1 MDGADALKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRCFPVCPYEHKRGQTQA 60
Db 1 MEGDADGKSGRRGGVQCICGDBGVGTAGDVFACDVCRCFPVCPYEHKRGQTQA 60
QY 61 CLOCKTKYKXHRGSPAIRGEGDDTDADGSDFNYPASGTEDQKOKIADRMRSWRMTGG 120
Db 61 CPQCKTKYKXHRGSPAIRGEGDDTDAD--SDFNYLASGNEQKOKIADRMRSWRMVGG 118
QY 121 SGNVGHPKYDSGETGLSKYDSGEIPRGVSPVTSQMSGEIPGASPDHMMSPGTNISR 180
Db 119 SGDVGRPKYDSGETGLTKYDSGEIPRGVSPVTSQISGEIPGASPDHMMSPGTNIGKR 178
QY 181 APFVYVNHSPNPSREFSGSICGNVAKERVDCGKWKQDKGATPMTNGTSIAPSEGAATDI 240
Db 179 APFVYVNHSPNPSREFSGSICGNVAKERVDCGKWKQDKGTIPMTNGTSIAPSEGVGDI 238
QY 241 DASTYNNMEDALLNDETRQPLSRKVPPIASSKINPYRMVIVLRLVVLSTIFLYRLTNPVRN 300
Db 239 DASTDYNMEDALLNDETRQPLSRKVPPLSSRINPYRMVIVLRLVVLSTIFLYRLTNPVRN 298
QY 301 AYPILWLLSVICETIWFALSWILDQPKPFPINRETYLDRLALRYDREGEPSQLAAVDIFVS 360
Db 299 AYPILWLLSVICETIWFALSWILDQPKPFPINRETYLDRLALRYDREGEPSQLAAVDIFVS 358
QY 361 TVDPLKEPPIVTANTVLSILAADVDPVKVSCVSDGASMLTFDALAETSEFARKWVPFV 420
Db 359 TVDPMKPPPLVTANTVLSILAADVDPVKVSCVSDGASMLTFDALAETSEFARKWVPFV 418
QY 421 KKYDIEPRAPDFYFCQIDYLDKQVQSFVDRAMKREYEEFKIRINALYSKALKVPEE 480
Db 419 KKYNIERAPDFYFQKIDYLDKQVHFSFVDRAMKREYEEFKRVNGLVAKAQKVPPE 478
QY 481 GWINQDGTWPNNTRHPGMIQVFLGHSGLDTEGNEPLRLVYVYREKRPFGOHKKAG 540
Db 479 GWINQDGTWPNNTRHPGMIQVFLGHSGLDTEGNEPLRLVYVYREKRPFGOHKKAG 538
QY 541 ANNALVRSVAVLTNGQVYMLNLDCHYINNSKAVREACFLMDNLPQVQVQRPQPDG 600
Db 539 ANNALVRSVAVLTNGQVYMLNLDCHYINNSKALREACFLMDNLPGRSVQVQRPQPDG 598
QY 601 IDRNDRYANRNVTFDINRLGLDGIQGPVYVGTGCVFNRTAIYGVPEPIKAKPGFLASL 660
Db 599 IDRNDRYANRNVTFDINRLGLDGIQGPVYVGTGCVFNRTAIYGVPEPIKAKPGFLSSL 658
QY 661 CGGKKKASKKRSDDKKSKNKHVDSSVPVNLEDIEEGVEGAGFPDDBKSVLMSQMSLEK 720
Db 659 CGGKKKASKK-K-GSDKKSKQKHVDSSVPVNLEDIEEGVEGAGFPDDBKSLMSQMSLEK 717
QY 721 RFGQSAFVASTLMEYGVGPOSSTPESILLKEAIIHVISCGYEDKSEMGTEIGWIVGSVTED 780
Db 718 RFGQSAFVASTLMEYGVGPOSATPESILLKEAIIHVISCGYEDKTEWGTGTEIGWIVGSVTED 777
QY 781 ILTGFKKHARGWRSVYCMKPEPAFKGSAPINLSDRLNQVLRWALGSVLEILSRHCPLMYG 840
Db 778 ILTGFKKHARGWRSIYCMKPEPAFKGSAPINLSDRLNQVLRWALGSVLEILSRHCPLMYG 837
QY 841 YGGRKLFERPAYINTIYPTSLPLLYCILPAICLLTGKFIEMPEISNLSIWFIALFL 900
Db 838 YGGRKLFERPAYINTIYPTSLPLLYCILPAICLLTGKFIEMPEISNLSIWFISLFI 897
QY 901 SIFATGILEMWSGVGIDEMWRNEQFVIGGISAHLEFVQGLLKLVLGADITNFTVTSKA 960
Db 898 SIFATGILEMWSGVGIDEMWRNEQFVIGGISAHLEFVQGLLKLVLGADITNFTVTSKA 957
QY 961 NDEEDGFAELYFKWTTLLIPIPTTILLINMVGVVAGTSYAINSGVQSGPLFGKLFPAFW 1020
Db 958 SDEEDGFAELYFKWTTLLIPIPTTILLINMVGVVAGTSYAINSGVQSGPLFGKLFPAFW 1017
QY 1021 VIVHLYPFLKGLMGQRNTPPTIVIVW 1046
Db 1018 VIVHLYPFLKGLMGQRNTPPTIVVW 1043

RESULT 5
AA184121
ID AA184121 standard; protein; 1076 AA.
XX
AC AA184121;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT Misc-difference 425
FT /note= "encoded by ANG"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
DR WPT; 2000-224343/19.
DR N-PSDB; AA299533.
XX
New genes which encode maize cellulose synthase polypeptides in plants
useful for modulating the expression of cellulose synthase in plants and
to produce transgenic plants expressing the novel protein.
XX
Claim 15; Page 196-199; 119pp; English.
XX
The present sequence represents a maize cellulose synthase polypeptide.
The cellulose synthase can be used for the improvement of stalk quality
for improved stand or silage. It also provides an increased concentration
of cellulose in the pericarp, hardening the kernel and improving its
handling ability. The sequences are used to produce transgenic plants and
seeds expressing the cellulose synthase. The polynucleotide is used for
modulating, preferably increasing, the level of the synthase in a plant
cell. The plants are preferably monocots. The polynucleotide is also used
as a probe or primer in the detection quantitation or isolation of gene
transcripts. The probes are useful in detecting deficiencies in the level
of mRNA in screenings for desired transgenic plant, for detecting or
mutations in the gene, for monitoring upregulation of compounds, for
changes in enzyme activity in screening assays of compounds, for
detection of any number of allelic variants of the gene, or for use as
molecular markers in plant breeding programs. The isolated nucleic acids
of the present invention can also be used for recombinant expression of
their encoded polypeptides or for use as immunogens in the preparation
and/or screening of antibodies. The proteins can be employed in assays
for enzyme agonists or antagonists of enzyme function or for use of
immunogens or antigens to obtain antibodies specifically immunoreactive
with a protein
XX
SQ Sequence 1076 AA;

Query Match 89.0%; Score 5143; DB 3; Length 1076;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 947; Conservative 69; Mismatches 59; Indels 6; Gaps 6;
QY 1 MDGADALKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRCFPVCPYEHKRGQTQA 60
Db 1 MDG-GDATNSGKHVAGVQCICGDBGVGTAGDGLFTACDVCRCFPVCPYEHKRGQTQA 59

PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			

Query Match

Best Local Similarity 78.4%; Score 4530; DB 3; Length 1054;

Matches 839; Conservative 98; Mismatches 104; Indels 26; Gaps 8;

Qy	19	QOICADGLGTLGDGVFTACDVCBPPVCRPCYEHKEGTQACLOCKTKYKRRHGS	PAIR 78
Db	9	QOICSDNVGKTVDGDRFVACDICSFPVCRPCYERKDGNSQPCQCKTRYKRLGSP	PAIP 68
Qy	79	GEEGDDTDADGS-DENYPASGTEDQOKIADRMRSMNTGGSGNVGHPKYDSGEIGLS	137
Db	69	GDKDEGLADEGTVFENYP-----QKEKISERMLGWHLTRGKEEMGEPOYDK	116
Qy	138	KYDSGEIPRGYVPSVTSNSQ-MSGEIPGASPDHMMSPGNTISRAPPFY-VNHS	PNPSRE 195
Db	117	-----EVSHNHLPLRTSRQDTSGEFSAASPERLSVSVSTIAGGKRLPYSSDVNQSP	NRRIV 171


```
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 03-JUL-1999; 99US-0142055P.
PR 04-JUL-1999; 99US-0142390P.
PR 05-JUL-1999; 99US-0142803P.
PR 06-JUL-1999; 99US-0142920P.
PR 07-JUL-1999; 99US-0142977P.
PR 08-JUL-1999; 99US-0143542P.
PR 09-JUL-1999; 99US-0143624P.
PR 10-JUL-1999; 99US-0144005P.
PR 11-JUL-1999; 99US-0144085P.
PR 12-JUL-1999; 99US-0144325P.
PR 13-JUL-1999; 99US-0144331P.
PR 14-JUL-1999; 99US-0144332P.
PR 15-JUL-1999; 99US-0144333P.
PR 16-JUL-1999; 99US-0144334P.
PR 17-JUL-1999; 99US-0144335P.
PR 18-JUL-1999; 99US-0144352P.
PR 19-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 22-JUL-1999; 99US-0145086P.
PR 23-JUL-1999; 99US-0145088P.
PR 24-JUL-1999; 99US-0145085P.
PR 25-JUL-1999; 99US-0145087P.
PR 26-JUL-1999; 99US-0145089P.
PR 27-JUL-1999; 99US-0145192P.
PR 28-JUL-1999; 99US-0145145P.
PR 29-JUL-1999; 99US-0145218P.
PR 30-JUL-1999; 99US-0145224P.
PR 31-JUL-1999; 99US-0145276P.
PR 01-AUG-1999; 99US-0145913P.
PR 02-AUG-1999; 99US-0145918P.
PR 03-AUG-1999; 99US-0145919P.
PR 04-AUG-1999; 99US-0145919P.
PR 05-AUG-1999; 99US-0145919P.
PR 06-AUG-1999; 99US-0146386P.
PR 07-AUG-1999; 99US-0146388P.
PR 08-AUG-1999; 99US-0146389P.
PR 09-AUG-1999; 99US-0147038P.
PR 10-AUG-1999; 99US-0147204P.
PR 11-AUG-1999; 99US-0147302P.
PR 12-AUG-1999; 99US-0147192P.
PR 13-AUG-1999; 99US-0147260P.
PR 14-AUG-1999; 99US-0147303P.
PR 15-AUG-1999; 99US-0147416P.
PR 16-AUG-1999; 99US-0147493P.
PR 17-AUG-1999; 99US-0147935P.
PR 18-AUG-1999; 99US-0148171P.
PR 19-AUG-1999; 99US-0148319P.
PR 20-AUG-1999; 99US-0148341P.
PR 21-AUG-1999; 99US-0148565P.
PR 22-AUG-1999; 99US-0148684P.
PR 23-AUG-1999; 99US-0149368P.
PR 24-AUG-1999; 99US-0149175P.
PR 25-AUG-1999; 99US-0149426P.
PR 26-AUG-1999; 99US-0149722P.
PR 27-AUG-1999; 99US-0149723P.
PR 28-AUG-1999; 99US-0149902P.
PR 29-AUG-1999; 99US-0149929P.
PR 30-AUG-1999; 99US-0150566P.
PR 31-AUG-1999; 99US-0150884P.
PR 01-SEP-1999; 99US-0151065P.
PR 02-SEP-1999; 99US-0151066P.
PR 03-SEP-1999; 99US-0151080P.
PR 04-SEP-1999; 99US-0151303P.
PR 05-SEP-1999; 99US-0151303P.
PR 06-SEP-1999; 99US-0151303P.
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.
PR 11-SEP-1999; 99US-0153758P.
PR 12-SEP-1999; 99US-0154018P.
PR 13-SEP-1999; 99US-0154039P.
PR 14-SEP-1999; 99US-0154779P.
PR 15-SEP-1999; 99US-0155139P.
PR 16-SEP-1999; 99US-0155486P.
PR 17-SEP-1999; 99US-0155659P.
PR 18-SEP-1999; 99US-0156458P.
PR 19-SEP-1999; 99US-0156596P.
PR 20-SEP-1999; 99US-0157117P.
PR 21-SEP-1999; 99US-0157753P.
PR 22-SEP-1999; 99US-0157865P.
PR 23-SEP-1999; 99US-0158029P.
PR 24-SEP-1999; 99US-0158232P.
PR 25-SEP-1999; 99US-0158369P.
PR 26-SEP-1999; 99US-0159233P.
PR 27-SEP-1999; 99US-0159294P.
PR 28-SEP-1999; 99US-0159295P.
PR 29-SEP-1999; 99US-0159329P.
PR 30-SEP-1999; 99US-0159330P.
PR 01-OCT-1999; 99US-0159331P.
PR 02-OCT-1999; 99US-0159637P.
PR 03-OCT-1999; 99US-0159638P.
PR 04-OCT-1999; 99US-0159584P.
PR 05-OCT-1999; 99US-0160741P.
PR 06-OCT-1999; 99US-0160767P.
PR 07-OCT-1999; 99US-0160768P.
PR 08-OCT-1999; 99US-0160770P.
PR 09-OCT-1999; 99US-0160814P.
PR 10-OCT-1999; 99US-0160815P.
PR 11-OCT-1999; 99US-0160815P.
PR 12-OCT-1999; 99US-0160980P.
PR 13-OCT-1999; 99US-0160981P.
PR 14-OCT-1999; 99US-0160989P.
PR 15-OCT-1999; 99US-0161040P.
PR 16-OCT-1999; 99US-0161405P.
PR 17-OCT-1999; 99US-0161406P.
PR 18-OCT-1999; 99US-0161359P.
PR 19-OCT-1999; 99US-0161360P.
PR 20-OCT-1999; 99US-0161361P.
PR 21-OCT-1999; 99US-0161920P.
PR 22-OCT-1999; 99US-0161922P.
PR 23-OCT-1999; 99US-0161993P.
PR 24-OCT-1999; 99US-0162142P.

Query Match 78.4%; Score 4530; DB 3; Length 1065;
Best Local Similarity 78.6%; Pred. No. 0;
Matches 839; Conservative 98; Mismatches 104; Indels 26; Gaps 8;

Qy 19 CQICADGLTTLGDGVFTACDVCRPCYHERKEGTQACLOCKTKYKRHRGSPAIR 78
Db 20 CQICSDNVGKTVDGRFVACDICSFPVCRPCYERKDGNCQCPQCKTRYKLGSPAIP 79

Qy 79 GEGDDTDADDGS-DFNYPASGTEDQKQIADRMRSWRMNTGGSGNVGHPKYDSEIGLS 137
Db 80 GKDDEDGLADEGTVFNYP-----QKEKISERMLGHLTRGKGEENGEPQYDK----- 127

Qy 138 KYDSGEIPRGVPSVTNSQ-MSGEIPGASPDHMMSTGNISRPAPPY-VNHSFNSRE 195
Db 128 -----EVSHNHLPRLTSRQDTSGEFSAASPERLSVSTIAGGKRLPYSSDVNQSPNRIV 182

Qy 196 FSGSIGNVAMKERVGDGWMKODKGAIPMTNGTSTAPSEGRAATDIDASTEXNMEDALLND 255
Db 183 DFGVGLNVAMKERVGDGWMKQKQNTGPV---STQASE-RGVVIDASTDILADEALLND 238

Qy 256 ETRQPLSRKVPFIASSKINPYRMVTLRLVLSIFLHYRLINPNVRNAYPLMLLSVCEIWF 315
Db 239 EARQPLSRKVISIPSSRINPYRMVIMLRVLVILCLFLHYRIINPNVFNALMLVSVCEIWF 298

Qy 316 ALSWILDQFPKFPFIPINRETYLDRLALYDREGPSQAAVDIFVSTVDPLKEPPIVTANT 375
Db 299 ALSWILDQFPKFPFVNRRETYLDRLALYDREGPSQAAVDIFVSTVDPLKEPPIVTANT 358
```

QY 376 VLSILAVDPVDKVCYSDDGASMLTFDALAETSEFARKVVPFVKYDIEBRAPEFYFC 435
Db VLSILAVDPVDKVCYSDDGASMLTFDALAETSEFARKVVPFVKYDIEBRAPEFYFC 418
QY 436 QKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALVPBEGWIMQDGTWPNGNNT 495
Db AKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALVPBEGWIMQDGTWPNGNNT 478
QY 496 RDHPGMIQVFLGHSGGLDTEGNEPRLVLYVVSREKPGFOHKKKAGAMNALVRVSALVTNG 555
Db RDHPGMIQVFLGHSGGLDTEGNEPRLVLYVVSREKPGFOHKKKAGAMNALVRVSALVTNG 538
QY 556 QYMLNLDCHYINNSKAVREACFLMDPNLGQVCVQVQPPQDFDGDIDRDRYANRNTVPF 615
Db PYMLNLDCHYINNSKALREACFLMDPNLGQVCVQVQPPQDFDGDIDRDRYANRNTVPF 598
QY 616 DINLRGLDGIQGVVVGTCVFNRTALVGYEPPIKA--KKPGFLASLCGGKKKASKKR 673
Db DINLRGLDGIQGVVVGTCVFNRTALVGYEPPIKVKKHKPSLLSKLGGSRKKNSKAKK 658
QY 674 SSDKKKSNKHVDSSVPVFNLEDEEGVEGAGDDEKSVLMSQMSLEKRFQSGAAFFVASTL 733
Db ESDKKKSGRHTDSTVPVFNLDDEEGVEGAGDDEKSVLMSQMSLEKRFQSGAAFFVASTL 718
QY 734 MEYGVPOSSTPESLLKEAIVHISCGYEDKSEWTEIGWYISGVTEDILTGFEMHARGWR 793
Db MENGVPSPSATENLLKEAIVHISCGYEDKSDWMEIGWYISGVTEDILTGFEMHARGWR 778
QY 794 SVYCKPKAPAFKSPAPINLSDRINQVLRWALGSVEILFSRHCPVWYGYGRILKFLERPAY 853
Db SYCKPKLPAFAKSPAPINLSDRINQVLRWALGSVEILFSRHCPVWYGYGRILKFLERPAY 838
QY 854 INTTLYPLTSLPLVVCILPAICLTGKFMPEINSLAGIWFIALPLSFATGILEMRWS 913
Db VNTIYPIISIFLLMCTLPVCLFTNQFIIFQISNIASIFLSLFSFATGILEMRWS 898
QY 914 GVGIDEMWNEQFWIGGISAHLFAVQGLLKVLGIDTFTVTSKANDEGDFAEIYMF 973
Db GVGIDEMWNEQFWIGGISAHLFAVQGLLKVLGIDTFTVTSKANDEGDFAEIYMF 958
QY 974 KWTLLIPIPTLLIINMGVAVGTSYAINSGVQSGPLFGKLFPAFWIIVHLYPFLKGLM 1033
Db KWTLLIPIPTLLIINMGVAVGTSYAINSGVQSGPLFGKLFPAFWIIVHLYPFLKGLM 1018
QY 1034 GRQNRTPTIIVWVLLASIFSLWVRVDPFTTRLAGPNIOTCGINC 1080
Db GRQNRTPTIIVWVLLASIFSLWVRVDPFTTRLAGPNIOTCGINC 1065

RESULT 8
ABB93356
ID ABB93356 standard; protein; 1065 AA.
XX AC ABB93356;
XX
DT 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 2567.
XX
KW Herbicidal; plant; agriculture; herbicide.
OS Arabidopsis thaliana.
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-BF009892.
XX
PR 28-AUG-2001; 2001WO-BF009892.
XX (FARB) BAYER AG.
PA
XX

PI Tietjen K, Weidner M;
XX WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX Claim 5; SEQ ID NO 2567; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90730-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 1065 AA;
Query Match 78.4%; Score 4530; DB 5; Length 1065;
Best Local Similarity 78.6%; Pred. No. 0;
Matches 839; Conservative 98; Mismatches 104; Indels 26; Gaps 8;
QY 19 COICADGLTLDGDVFTACDVCREFVCPYCHERKECTOACLOKTKYKHRGSPAIR 78
Db 20 COICSDNVKTVGDGRFVACDICSPFVCPYCHERKECTOACLOKTKYKHRGSPAIR 79
QY 79 GEGEDDTDADGGS-DPNYPASGTEQOKIADRMESWRMNTGSGNVGHPKPYDSGEIGLS 137
Db 80 GDXDEGLADEGTVEFNYP-----QKEKISERMLGWLHTRGKGEEMGEFQYDK----- 127
QY 138 KYDSGEIPRGYVPSVTNSQ-MSGEIPGASPDHMHMSPTGNISRRAPFPY-VNHSPNPSRE 195
Db 128 -----EVSHNHLPLRTSRQDTSGETSAASPERLSVSTIAGGKRLFPSSDVNQSPNRIV 182
QY 196 FSGSITGNVAKERVDMKQKQKGAIPMTNGTISIAPSEGRAATDIDASTEYNMEDALLND 255
Db 183 DPGVGNVAKERVDMKQKQKGAIPMTNGTISIAPSEGRAATDIDASTEYNMEDALLND 238
QY 256 ETROPLSRKVPITASSKINPYRMVIVLRLVSLIFLHYRLTNVPRNAYPLWLLSVICEIWF 315
Db 239 EARQPLSRKVSIPSSRIINPYRMVIMRLVLCLFLHYRLTNVPRNAYPLWLLSVICEIWF 298
QY 316 ALSWILDQFPKWPPIINRETYLDRALRYDREGEPSQLAAVDIFVSTVDPLEKPPVIVTANT 375
Db 299 ALSWILDQFPKWPPIINRETYLDRALRYDREGEPSQLAAVDIFVSTVDPLEKPPVIVTANT 358
QY 376 VLSILAVDPVDKVCYSDDGASMLTFDALAETSEFARKVVPFVKYDIEBRAPEFYFC 435
Db 359 VLSILAVDPVDKVCYSDDGASMLTFDALAETSEFARKVVPFVKYDIEBRAPEFYFC 418
QY 436 QKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALVPBEGWIMQDGTWPNGNNT 495
Db 419 AKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALVPBEGWIMQDGTWPNGNNT 478
QY 496 RDHPGMIQVFLGHSGGLDTEGNEPRLVLYVVSREKPGFOHKKKAGAMNALVRVSALVTNG 555
Db 479 RDHPGMIQVFLGHSGGLDTEGNEPRLVLYVVSREKPGFOHKKKAGAMNALVRVSALVTNG 538
QY 556 QYMLNLDCHYINNSKAVREACFLMDPNLGQVCVQVQPPQDFDGDIDRDRYANRNTVPF 615
Db 539 PYMLNLDCHYINNSKALREACFLMDPNLGQVCVQVQPPQDFDGDIDRDRYANRNTVPF 598
QY 616 DINLRGLDGIQGVVVGTCVFNRTALVGYEPPIKA--KKPGFLASLCGGKKKASKKR 673
Db 599 DINLRGLDGIQGVVVGTCVFNRTALVGYEPPIKVKKHKPSLLSKLGGSRKKNSKAKK 658
QY 674 SSDKKKSNKHVDSSVPVFNLEDEEGVEGAGDDEKSVLMSQMSLEKRFQSGAAFFVASTL 733

Db 659 ESDKKSGRHTDSTVPVFNLDDEEGVEGAGFDDDEKALLMSQMSLEKRFQGSVAVFVASTL 718
QY 734 MEYGVQDSTPESLLKEAHVHISCGYEDKSEWTEIGWYGSVTEIDLTGFKMHARGWR 793
Db 719 MENGVPSPATPENLLKEAHVHISCGYEDKSDWGMGEIGWYGSVTEIDLTGFKMHARGWR 778
QY 794 SVYCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGRKLFERFAY 853
Db 779 SIYCMKPLPAKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGRKLFERFAY 838
QY 854 INTTIVPLTSLPVLVYICILPAICLLTGKFINPEISNLASIFIALFISFATGILEMRWS 913
Db 839 VNTTIVPITSIPILMYCTLPVCLFTNQFIIPQISNIASIFLSLFSIFATGILEMRWS 898
QY 914 GVGDEWRNEQFVIGGSAHLFAVQGLLKVLGADTNTFTVTSKANDEGDPFAELMYF 973
Db 899 GVGDEWRNEQFVIGGSAHLFAVQGLLKVLGADTNTFTVTSKANDEGDPFAELMYF 958
QY 974 KWTLLIPPTLLIINMVGVSAGTSYAINSGYQSGPLFGKLPFAFWIVHLYPFLKGLM 1033
Db 959 KWTLLIPPTLLIINMVGVSAGTSYAINSGYQSGPLFGKLPFAFWIVHLYPFLKGLM 1018
QY 1034 GRQNRTPVIVWVLLASIFSLWVRVDPPTTLRAGNIQTGCINC 1080
Db 1019 GRQNRTPVIVWVLLASIFSLWVRVDPPTTLRAGNIQTGCINC 1065

RESULT 9
AAW33819
ID AAW33819 standard; protein; 1065 AA.
XX
AC AAW33819;
XX
XZ 17-OCT-2003 (revised)
DT 06-JUN-1998 (first entry)
XX
DE Arabidopsis cellulose biosynthetic protein Ath-B.
XX
XW Cellulose; cellulose synthase; RSW1 protein; beta-1,4-glucan;
KW transgenic plant.
XX
OS Arabidopsis thaliana; var. Columbia.
XX
PN WO9800549-A1.
XX
XZ 08-JAN-1998.
XX
XX 24-JUN-1997; 97WO-AU000402.
XX
XX 27-JUN-1996; 96AU-00000699.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
PI Arioli A, Williamson RE, Betzner AS, Peng L;
XX
DR WPI; 1998-086974/08.
DR N-PSDB; AAW06567.
XX

DNA encoding cellulose biosynthetic enzyme - useful for manipulation of
cellulose and beta-1,4-glucan.
XX
PS Claim 29; Page 144-149; 207pp; English.
XX
CC This polypeptide is encoded by cDNA clone Ath-B (see AAW06567) that was
CC isolated from an Arabidopsis thaliana cDNA library using PCR primers (see
CC AA79632-34) based on cellulose synthase RSW1 genomic clone 23H12 (see
CC AAW06563) and EST clone AAT20782 (see AAW06562). It is closely related to
CC Arabidopsis cellulose synthase RSW1 protein (see AAW33816-17). Claimed
CC nucleic acid molecules (see AAW06562-69) coding for claimed polypeptides
CC (see AAW33816-20 and AAW46202) involved in cellulose biosynthesis can be
CC used to manipulate the cellulose and/or beta-glucan content of transgenic
CC plants. Expression of nucleic acids in the sense orientation increases

CC the level of cellulose and reduces the level of non-crystalline beta-1,4-
CC glucan and starch, providing plants with modified strength and/or shape
CC and/or fibre properties, or having increased resistance to stresses or
CC pests. Antisense, ribozyme or co-suppression molecules can be used to
CC reduce the cellulose content of a transgenic plant, e.g. to improve
CC digestibility or to alter carbon partitioning such that increased carbon
CC is available for growth, rather than deposited as cellulose. (Updated on
CC 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1065 AA;
Query Match 77.9%; Score 4499; DB 2; Length 1065;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 835; Conservative 98; Mismatches 108; Indels 26; Gaps 8;
QY 19 CQICADGLGTLTLDGVFTACDVCPVPCYERKEGTQACLOCKTKYKRRHGSFAIR 78
Db 20 CQICSDNVGTVGDRFVACDICSFPVPCYERKDGNSQCPCKTRVYKRLKGSFAIP 79
QY 79 GEGDDTDADDGS--DFNYPASGTEDOKIADRMRSRMNTGGSGNVGHPKYDSGEIGLS 137
Db 80 GKDDEGLADEGTVEFNYP-----QKEKISERMLGWLHTRGKEEMGEPOYDK----- 127
QY 138 KYDSGEIPIRGVPSVTSNQ-MSGEIPGASPDHMHMSPTGNISRAPPY-VNHSPNFSRE 195
Db 128 -----EVSHNHLPLRTSRQDTSGEFSAASPERLSVSTIAGGKLPYSSDVNQSPNRV 182
QY 196 FSGSIGNVANKERVDMKQDKGAIPTWNGTSTAPSEGRANTDIDASTEYVNMEDALND 255
Db 183 DPVGLGNVANKERVDMKQDKGAIPTWNGTSTAPSEGRANTDIDASTEYVNMEDALND 238
QY 256 ETROPLSRKVIASSKINPYRMVILRLVLSIFLHYRLTNPNVAYPLWLLSVICETWF 315
Db 239 EAROLLSRKVISIPSSRINPYRMVIMLRVLILCLPLHYRITNPVFNALMLVSVICETWF 298
QY 316 ALSWILDQFPKWFPIINRETYLDRALRYDREGESQLAAVDIFVSTVDPLKEPPIVANT 375
Db 299 ALSWILDQFPKWFPIINRETYLDRALRYDREGESQLAAVDIFVSTVDPLKEPPIVANT 358
QY 376 VLSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPVKYKIDIEPRAPPEYFC 435
Db 359 VLSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPVKYKIDIEPRAPPEYFC 418
QY 436 QKIDYLDKQVPSFKDRBRAMKYEFEFKIRINALVSKALKVPEEGIMODGTPKGNNT 495
Db 419 AKIDYLDKQVPSFKDRBRAMKYEFEFKIRINALVSKALKVPEEGIMODGTPKGNNT 478
QY 496 RDHPGMIQVFLHSGGLDTEGNELPLVYVSREKRPQFHHKKAGAMNALVRVSAVLTING 555
Db 479 GDHGMIOVFLGQNGGLDAEGNELPLVYVSREKRPQFHHKKAGAMNALVRVSAVLTING 538
QY 556 QYMLNDCDHYNNKSKAVREAMCFMDPNLGPQVYQVQFQFQFQFQFQFQFQFQFQFQF 615
Db 539 PFILNDCDHYNNKSKALREAMCFMDPNLGLKQVYQVQFQFQFQFQFQFQFQFQFQFQF 598
QY 616 DINLRGLDGTQGVYVGTGCVFNRTAIVGYEPPITKA--KKPGFLASLCGGKAKSKSKR 673
Db 599 DINLRGLDGTQGVYVGTGCVFNRTAIVGYEPPITKA--KKPGFLASLCGGKAKSKSKR 658
QY 674 SSDKKKSNKHVDSSVPVFNLEIDIEEGVEGAGFDDDEKALLMSQMSLEKRFQGSVAVFVASTL 733
Db 659 ESDKKSGRHTDSTVPVFNLDDEEGVEGAGFDDDEKALLMSQMSLEKRFQGSVAVFVASTL 718
QY 734 MEYGVQDSTPESLLKEAHVHISCGYEDKSEWTEIGWYGSVTEIDLTGFKMHARGWR 793
Db 719 MENGVPSPATPENLLKEAHVHISCGYEDKSDWGMGEIGWYGSVTEIDLTGFKMHARGWR 778
QY 794 SVYCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGRKLFERFAY 853
Db 779 SIYCMKPLPAKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGRKLFERFAY 838
QY 854 INTTIVPLTSLPVLVYICILPAICLLTGKFINPEISNLASIFIALFISFATGILEMRWS 913

DT 03-JUL-2000 (first entry)
 XX Amino acid sequence of a maize cellulose synthase.
 DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 XX transgenic plant; plant breeding marker.
 KW Zea mays.
 OS
 XX W0200009706-A2.
 PN 24-FEB-2000.
 XX 16-AUG-1999; 99WO-US018760.
 XX 17-AUG-1998; 98US-0096822P.
 PR (PION-) PIONEER HI-BRED INT INC.
 XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX WPI; 2000-224343/19.
 DR N-PSDB; AAZ99506.
 XX New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.
 XX Claim 15; Page 126-128; 119pp; English.
 XX The present sequence represents a maize cellulose synthase polypeptide.
 CC The cellulose synthase can be used for the improvement of stalk quality
 CC for improved stand or silage. It also provides an increased concentration
 CC of cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for use as
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 XX SQ Sequence 1074 AA;
 Query Match 69.1%; Score 3993.5; DB 3; Length 1074;
 Best Local Similarity 68.4%; Pred. No. 0;
 Matches 749; Conservative 134; Mismatches 153; Indels 59; Gaps 20;
 QY 2 DGDADA-LKSGRHGAGVQCQICDGLGTTLDGVFTACDVCRFPVCPYEHKEGTQA 60
 DB 23 DGDAPVPAKPTKSANGQVCQICGTVGVSATGDFVACNECAFPVCPYEHKEGNQC 82
 QY 61 CLOCTKYKRRGSPATRGEGDDTDADD-GSDFNYPASGTEDOKOKIADMRSRWNTG 119
 DB 83 CPQCKTRYKQKSGSPRVHGDD-EEDVDLLEFNFY-----KQ----- 119
 QY 120 GSGNVGHPKY----DSGEIGLS-KYDSGEIPRGVPSVTN-SOMSGEIPGASPDHH-MMS 172
 DB 120 --GNKGPEWQLQGDADLSSARHD-----PHHRIPLTSGQISGIPDASPDHRSIR 173
 QY 173 PTGN-ISRRAPFPYVNHSPNPSREF-SGSIGNVAWKRVDGKWKQKGAIPMTNGTSIA 230
 DB 174 PTSSYVDPSPVPV--RIVDFSKDLSNGLSVDKERVESVRVKQDKNMLQVTN----- 226

QY 231 PSEGRAATDIDASTEYNMEDALLNDETROPISRKVPPIASSKINPYRMVIVILVLSIFL 290
 DB 227 -KYPEARGDME-GTGSNGEDMQWVDARLPLSRITVPISSNQLNRIYVILRLILFFF 284
 QY 291 HYRLTNVRNAYPLWLLSVLCEIWFALSWLLDOPPKWFFPINRETYLDRALRYDREGEPS 350
 DB 285 QYRISHVRNAYGLNLVSVICEVWFALSWLLDQPKWYPINRETYLDRALRYDREGEPS 344
 QY 351 QLAADVIFVSTVDPLKEPPIVNTANTVLSILAVDYPVDKVCYVSDDGASMLTFDALETS 410
 DB 345 QLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVCYVSDDGASMLTPESLSETA 404
 QY 411 EFARKWVPFVKKYVDIEPRAPEFFCOKIDYLKQVQSFVKDRAMKREYEEFKIRINAL 470
 DB 405 EFARKWVPFCKKHNIIEPRAPEFFAOKIDYLDKQISFVKERRAMKREYEEFKIRINAL 464
 QY 471 VSKALVPEEGWTIMQDGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVVSREKR 530
 DB 465 VAKAQVPEEGWTMADGTAMPNPRDHPGMIQVFLGHSGLDTEGNEPLRLVVSREKR 524
 QY 531 PGFOHHKAGAMNALVRVSAVLINGQYMLNLDCHYINNKAVERAMCFMDPMLGPQVC 590
 DB 525 PGFOHHKAGAMNALIRVSAVLINGAYLLNVDCDHYFNSSKALREAMCFMMDPALGRKTC 584
 QY 591 YVOPFORFDGIDRNDRYANENTVFFDINRLDGIQGPVVYVGTGCVENRTAIYGEVEPIK 650
 DB 585 YVOPFORFDGIDLHDYRANRNI VFFDINMGLDGIQGPVVYVGTGCCFNQALYDVPVT 644
 QY 651 AK--KPGFLASLGGKKKASKKRSDDKKSNKHVDSSVPFVNLEIDIEEGVEGAGFDE 708
 DB 645 EADLEPNIVKSCGRK-RKNKSYMDSQSRIMKRTESSAPIFNWEDIEEGIE--GYEDE 701
 QY 709 KSVLMSQMSLEKPGSOAAVASTLMYEGVGVQSSSTPESLLKEAHHVISCGYEDKSEWGT 768
 DB 702 RSVLMSQRKLEKRFQSPPIASTFTMTQGGIPSTNPASLLKEAHHVISCGYEDKTEWGW 761
 QY 769 EIGWYGSVTEIDLTGFMHARGWSVYCMKPAFKGSAPINLSRDLNOVLRWALGSVE 828
 DB 762 EIGWYGSVTEIDLTGFMHARGWSQSIYCMPPRPFCKGSAPINLSRDLNOVLRWALGSVE 821
 QY 829 ILSRHCPLWVGGRKLFLERFAYINTTYPLTSLPLLVVLCILPAICLTGKIMPEIS 888
 DB 822 ILSRHCPIWYNGRKLKLERLAYINTIVYPTSVPLIAYCVLPALCLLTNKEIIEIS 881
 QY 889 NLSIMPIALFLSIPATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFVAVFQGLLKVLA 948
 DB 882 NYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWIGGTSAHLFVAVFQGLLKVLA 941
 QY 949 GIDTNFTVTSKANDEGDFAEIYMFKWTLLIPPTTILINMWGVAGTSVAINSGYQSW 1008
 DB 942 GIDTNFTVTSKANDEGDFAEIYVFKWTSLLIPPTTVILNMGVAGISVAINSGYQSW 1001
 QY 1009 GPLFGKLFPAFWIVHLYPFLKGLMGQRNTPITIVWAVLLASIFSLWVRVDPF--T 1065
 DB 1002 GPLFGKLFPSIWIHLHLYPFLKGLMGQRNTPITIVWISILLASIFSLWVKIDPFISPT 1061
 QY 1066 TRLAGPNIOTCGNC 1080
 DB 1062 QKAAA--LGQCGVNC 1074
 RESULT 13
 AAY84117
 ID AAY84117 standard; protein; 1074 AA.
 XX AC AAY84117;
 XX XX
 DT 03-JUL-2000 (first entry)
 DE Amino acid sequence of a maize cellulose synthase.
 XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 17:19:28 ; Search time 172 Seconds

(without alignments)
3484.575 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVCO.....VDPPTTLAGPNIQTGCING 1080

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09900237/runat_18082004_081515_14658/app_query.fasta_1.1223
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09900237 @cgn_1_84 @runat_18082004_081515_14658 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCUTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	4499	77.9	3614	4	US-09-221-013A-9
2	3963.5	68.6	3603	4	US-09-221-013A-5
3	3959.5	68.5	3673	4	US-09-221-013A-11
4	3633.5	62.9	3828	4	US-09-221-013A-7
5	3518.5	60.9	8411	4	US-09-221-013A-3
6	3477	60.2	3328	3	US-08-960-048-1
7	3477	60.2	3328	4	US-09-838-586-1
8	2629.5	45.5	2248	4	US-09-221-013A-1
9	2539.5	44.0	5009	4	US-09-221-013A-4
10	1833	31.7	4612	3	US-08-960-048-2
11	1833	31.7	4612	4	US-09-838-586-2
12	1756.5	30.4	1741	4	US-09-221-013A-13

13	478	8.3	281	4	US-09-313-294A-6485	Sequence 6485, Ap
14	477	8.3	300	4	US-09-313-294A-4753	Sequence 4753, Ap
15	442	7.6	291	4	US-09-313-294A-5519	Sequence 5519, Ap
16	401.5	6.9	268	4	US-09-313-294A-849	Sequence 849, App
17	389.5	6.7	286	4	US-09-313-294A-4696	Sequence 4696, Ap
18	349	6.0	284	4	US-09-313-294A-3993	Sequence 3993, Ap
19	312	5.4	287	4	US-09-313-294A-5274	Sequence 5274, Ap
20	283	4.9	283	4	US-09-313-294A-1981	Sequence 1981, Ap
21	279	4.8	278	4	US-09-313-294A-4015	Sequence 4015, Ap
22	265	4.6	317	4	US-09-313-294A-7123	Sequence 7123, Ap
23	251.5	4.4	2841	4	US-09-489-039A-3885	Sequence 3885, Ap
24	246	4.3	9540	1	US-07-689-008-1	Sequence 1, Appli
25	241	4.2	159	4	US-09-313-294A-1693	Sequence 1693, Ap
26	241	4.2	16836	4	US-09-147-236-1	Sequence 1, Appli
27	241	4.2	16836	4	US-09-147-236-10	Sequence 10, Appli
28	241	4.2	16836	4	US-09-522-474-1	Sequence 1, Appli
29	241	4.2	16836	4	US-09-522-474-10	Sequence 10, Appli
30	238.5	4.1	2262	1	US-07-689-008-3	Sequence 3, Appli
31	236.5	4.1	2217	4	US-09-543-681A-3356	Sequence 3356, Ap
32	230	4.0	2652	4	US-09-543-681A-2775	Sequence 2775, Ap
33	210.5	3.6	2112	4	US-09-489-039A-3649	Sequence 3649, Ap
34	209	3.6	2912	4	US-08-541-939-1	Sequence 1, Appli
35	209	3.6	2912	5	FCT-US91-01726-2	Sequence 2, Appli
36	195.5	3.4	1707	4	US-09-489-039A-3798	Sequence 3798, Ap
37	192.5	3.3	1863	4	US-09-489-039A-518	Sequence 518, App
38	189.5	3.3	250	4	US-09-313-294A-582	Sequence 582, App
39	136.5	2.4	2673	4	US-09-252-991A-2993	Sequence 2993, Ap
40	136.5	2.4	2715	4	US-09-252-991A-2686	Sequence 2686, Ap
41	136.5	2.4	2814	4	US-09-252-991A-2875	Sequence 2875, Ap
42	134	2.3	7430	4	US-08-956-171E-260	Sequence 260, App
43	133	2.3	441529	3	US-09-103-840A-1	Sequence 1, Appli
44	131	2.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
45	128	2.2	2946	4	US-09-252-991A-7660	Sequence 7660, Ap

ALIGNMENTS

RESULT 1
US-09-221-013A-9
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Bettner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(3411)
US-09-221-013A-9

Alignment Scores:
Pred No.: 0
Score: 4499.00
Percent Similarity: 87.44%
Best Local Similarity: 78.26%
Query Match: 77.86%
DB: 4
Length: 3614
Matches: 835
Conservative: 98
Mismatches: 108
Indels: 26
Gaps: 8

US-09-900-237A-30 (1-1080) x US-09-221-013A-9 (1-3614)

Qy	19	CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys	38
Db	274	TGCCAGATCTGTAGTCAATTTGGCAAGACTGTTGATGAGATCGTTTGTGGCTGT	333
Qy	39	AspValCysArgPheProValCysArgProCysIleTyrGluHisGluArgLysGluGlyThr	58
Db	334	GATATTTGTTCAITCCCAAGTTTGGCGCTTGTACGAGTATGAGAGGAAGATGGGAAT	393
Qy	59	GlnAlaCysLeuGlnCysLysThrLysIleArgHisArgGlySerProAlaIleArg	78
Db	394	CAATCTTGCTCGTCAAGTCAAGAGGCTCAAGAGGTAGTCTGCTGATTTCT	453
Qy	79	GlyGluGluGlyAspAspThrAspAlaAspAspGlySer---AspPheAsnTyrProAla	97
Db	454	GGTATTAAGACGAGGATGGCTTAGCTGATGAAGGTACTGTGTGATTCAACTACCTC---	510
Qy	98	SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrPargMetAsn	117
Db	511	-----CAGAAGGAGAAATTTACAGACGGATGCTTGGTTGGCATCTTTACT	555
Qy	118	ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer	137
Db	556	CTGGGAAGGAGAGGAATGGGGAACCCAGATATGATAA-----	597
Qy	138	LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln---	156
Db	598	-----GAGGCTCTCAACATCATCTTCTCGTCTCAGCAGCAGACAGAT	642
Qy	157	MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn	176
Db	643	ACTTCAGGAGAGTTTTCGTGCTCCTACCTGAACGCTCTGTGTATCTTCTACTATCGCT	702
Qy	177	IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu	195
Db	703	GGGGAAAGCGCTTCCCTATTCTACAGATGTCATCAATCACCAAATGAAGAGATGTG	762
Qy	196	PheSerGlySerIleGlyAsnValAlaTyrLysGluArgValAspGlyTyrPlysMetLys	215
Db	763	GATCCTGTGGACTCGGAATGTAGCTTGGAGAGAGAGTTGATGCTGGAAATGAAG	822
Qy	216	GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg	235
Db	823	CAAGAGAAGAATACTGGTCTCTGTC-----AGCAGCAGGCTGTTCTTGAA---AGA	870
Qy	236	AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp	255
Db	871	GGTGGATGATATTTGATGCAGCAGACAGATATCTCAGAGATGAGGCTCTGCTGAATGAC	930
Qy	256	GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr	275
Db	931	GAAGCGAGCGAGCTTCTGTCAAGGAAAGTTTCAATTCCTTCATCACGGATCAATCTCTTAC	990
Qy	276	ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr	295
Db	991	AGAATGGTTATATGCTGCGGCTTGTATCCTTTGTCTTCTTGTGCATTAACCGTAAACA	1050
Qy	296	AsnProValArgAsnAlaTyrProLeuTyrLeuLeuSerValIleCysGluIleTyrPhe	315
Db	1051	AAACCAGTGCCAAATGCTTTGCTCTATGGCTGGTCTCTGTGATATGTGAGATCTGGTTT	1110
Qy	316	AlaLeuSerThrIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyr	335
Db	1111	GCCTTATCTCGATTTTGGATCAGTTTCCCAAGTGGTTTCTCTGTGAAACCGTGAACCTAC	1170
Qy	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaIleVal	355
Db	1171	CTCGACAGCTTGTCTTTAAGATATGATCGTGAAGGTGAGCCATCACAGTTAGCTGCTGTT	1230
Qy	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
Db	1231	GACATTTTCGTGAGTACTGTTGACCCCTTGAAGAGAGCACCCTTGTGACAGCCAAACA	1290

Qy	376	Val	Leu	Ser	Ile	Leu	Ala	Val	Asp	Yr	Pro	Val	Asp	Val	Ser	Cys	Tyr	Val	Ser	Asp	395
Db	1291	GTG	CTC	TCT	AT	ATC	TGG	CGT	TG	TG	ACT	ACC	CAG	TG	TCT	GT	TAT	GT	TG	TCT	1350
Qy	396	Asp	Gly	Ala	Ser	Met	Leu	Thr	Phe	Asp	Ala	Leu	Ala	Glu	Thr	Ser	Glu	Phe	Ala	Arg	415
Db	1351	GAT	GGT	GCT	GCT	AT	TG	TAT	CAT	TG	TAAT	CAC	TTC	GCA	AAA	CAT	CAG	AGT	TTG	TCT	1410
Qy	416	Trp	Val	Pro	Phe	Val	Lys	Lys	Tyr	Asp	Ile	Glu	Pro	Arg	Ala	Pro	Glu	Phe	Tyr	Phe	435
Db	1411	TGG	GT	TAC	CAAT	TTT	GCA	AAA	AAT	TAG	CAT	TAG	AGC	CT	CGT	CAC	CAG	AAT	TG	TCT	1470
Qy	436	Gln	Lys	Ile	Asp	Tyr	Ile	Lys	Asp	Lys	Val	Gln	Pro	Ser	Phe	Val	Lys	Asp	Arg	Ala	455
Db	1471	GCG	AAA	AAT	TAG	AT	CTT	GAA	GGA	TAA	AGT	TTC	AG	CAT	CAT	TTG	TG	CAA	AGAT	CG	1530
Qy	456	Met	Lys	Arg	Glu	Tyr	Glu	Glu	Phe	Lys	Ile	Arg	Ile	Asn	Ala	Leu	Val	Ser	Lys	Ala	475
Db	1531	ATG	AAG	AGG	GAAT	TAT	GAG	GAAT	TTAA	AAAT	CCG	AAT	CAAT	CAAT	CAAT	CAAT	CAAT	CAAT	CAAT	CAAT	1590
Qy	476	Lys	Val	Pro	Glu	Glu	Glu	Gly	Tyr	Trp	Ile	Met	Gln	Asp	Gly	Thr	Pro	Trp	Pro	Gly	495
Db	1591	AAAT	TG	TCT	CT	GA	AAG	AGG	TGG	TG	TAT	TCA	CAG	ATG	GC	CAC	CGT	GGC	CT	GAAA	1650
Qy	496	Arg	Asp	His	Pro	Gly	Met	Ile	Gln	Val	Phe	Leu	Gly	His	Ser	Gly	Gly	Leu	Asp	Thr	515
Db	1651	GGG	GAC	CAAT	CA	CGA	AT	GAT	CC	AGG	TCT	CTT	TAG	GGC	AAA	AT	TGT	GAT	CTT	GAT	1710
Qy	516	Gly	Asn	Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro	Gly	Phe	Gln	535
Db	1711	GGC	AAT	CAG	CT	CCG	CGT	TTT	GAT	TAT	TGT	TTCT	CGA	AAA	AGG	CAC	CAG	AT	TCC	CAG	1770
Qy	536	His	Lys	Lys	Ala	Gly	Ala	Met	Asn	Ala	Leu	Val	Arg	Val	Ser	Ala	Val	Leu	Thr	Asn	555
Db	1771	CAC	AAA	AA	GGC	TGT	GT	CT	AT	GAAT	GC	ACT	GGT	GAG	AGT	TTT	CAG	CAG	TTCT	TAC	1830
Qy	556	Gln	Tyr	Met	Leu	Asn	Leu	Asp	Cys	Asp	His	Tyr	Ile	Asn	Asn	Ser	Lys	Ala	Val	Arg	575
Db	1831	CCT	TT	CA	CT	TT	GAAT	CT	TG	AT	CT	TAT	TAC	ATA	ATA	TAAC	CA	CA	AGC	CT	1890
Qy	576	Ala	Met	Cys	Phe	Leu	Met	Asp	Pro	Asn	Leu	Gly	Pro	Gln	Val	Cys	Tyr	Val	Gln	Phe	595
Db	1891	GCA	AT	TG	CT	TT	CT	CT	CT	GA	CGA	CC	CA	AA	CT	CGG	GA	AG	CA	AGT	1950
Qy	596	Gln	Arg	Phe	Asp	Gly	Ile	Asp	Arg	Asn	Asp	Arg	Tyr	Ala	Asn	Arg	Asn	Thr	Val	Phe	615
Db	1951	CAA	AGAT	T	T	GAT	G	TAT	CG	ATA	GA	CA	TAG	AT	GCT	TAAT	CT	GA	TAT	AC	2010
Qy	616	Asp	Ile	Asn	Leu	Arg	Gly	Leu	Asp	Gly	Ile	Gln	Gly	Pro	Val	Tyr	Val	Gly	Thr	Gly	635
Db	2011	GAT	AT	TAA	CT	T	GAG	AGG	TTT	AG	TAT	G	GG	AT	T	CA	AG	CA	CT	GG	2070
Qy	636	Val	Phe	Asn	Arg	Thr	Ala	Ile	Tyr	Gly	Tyr	Glu	Pro	Pro	Ile	Lys	Ala	-----	Lys	Lys	653
Db	2071	GTT	TT	TCA	CA	GA	AA	CA	CA	GAT	TAT	CGG									

```

QY 734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle 753
Db 2371 ATGGAATAATGGTGGTGTCTCTCTCAGCAACATCCAGAAACATTTCTCAAGAGGCTATC 2430
QY 754 HisValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIle 773
Db 2431 CATGTCTATTAGTTGTGGTATAGGATAAGTCAGATTGGGGAATGGAGATTGGATGCATC 2490
QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArg 793
Db 2491 TATGGTTCTGTGACAGAAGATATTCTGACTGGGTTCAAAATGCGATGCCCGTGGATGCCGA 2550
QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
Db 2551 TCCATTACTTGCATGGCTTAAGCTTCCAGCTTTTCAAGGGTTCGTCTCTATCAATCTTTCA 2610
QY 814 AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg 833
Db 2611 GATCGTCTGAACCAAGTGTGAGTGGCTTTAGGTTTCAGTTGAGATTCTTTCAGTCCG 2670
QY 834 HisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
Db 2671 CATTTGCTATATGATGGTTTCAATGGGAGGCTAAATTTCTTGAGAGGTTTGGCTAT 2730
QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuPro 873
Db 2731 GTGAACACCAACCATCTACCTATCACTTCCATCTCTTCTCAATGATTGTACATTGCTA 2790
QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluLysSerAsnLeuAlaSerile 893
Db 2791 GCCGTTGTCTCTTCCAAACCAAGTTATTATTATCTCAGATTAGTAACATTGCCAAGTATA 2850
QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer 913
Db 2851 TGGTTTCTGCTCTCTCTCTCTCTCCATTTCCGCCAGCGGTATCTAGAAATGAGGTGGAGT 2910
QY 914 GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSer 933
Db 2911 GGCCTAGGCAATACAGAAATGGTGAGAAACGAGCAGTTTGGGTTCATTGGTGAGTAACC 2970
QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
Db 2971 GCTCATTTATTCGTGTGTTTCAAGGTATCTCTCAAGTCTCTTCCGCGTATTGACACAAC 3030
QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
Db 3031 TTCACAGTTTACCTCAAAAGCTTCAGATGAAGACGAGACTTTGCTGAGCTCTACTTCTTC 3090
QY 974 LysTrpThrThrLeuLeuProProThrThrIleLeuIleAsnMetValGlyVal 993
Db 3091 AAATGGACAACACTTCTGATTCGCCCAACGAGCTGCTCATTTGTAACCTTAGTGGGAGTT 3150
QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013
Db 3151 GTTCGAGGAGTCTCTATGATCTCAACAGTGGATACCAATCATGGGACCACTCTTTGGT 3210
QY 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
Db 3211 AAGTTGTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3270
QY 1034 GlyArgGlnAsnArgTrpProThrIleValIleValTrpAlaValLeuAlaSerile 1053
Db 3271 GGTGAGAGAACCGGATCTCTACCATTTGTTGGTCTGGTCTGCTGCTGCTGCTGCTGCTGCT 3330
QY 1054 PheSerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
Db 3331 TTCTCGTTGTGGTTAGATTGATCCCTTCTACAGCGAGTCACTGGCCCGGACATT 3390
QY 1074 GlnThrCysGlyIleAsnCys 1080
Db 3391 CTGGAATGTGGAATCAACTGT 3411

```

RESULT 2

```

US-09-221-013A-5
; Sequence 5, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3603
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3243)
US-09-221-013A-5
Alignment Scores:
Pred. No.: 0 Length: 3603
Score: 3963.50 Matches: 733
Percent Similarity: 80.80% Conservatives: 138
Best Local Similarity: 68.00% Mismatches: 168
Query Match: 68.60% Indels: 39
DB: 4 Gaps: 15
US-09-900-237A-30 (1-1080) x US-09-221-013A-5 (1-3603)
QY 2 AspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGluIle 21
Db 73 GATGGCGGGACCAACCTTTTGAAGATATGAAT-----GGCCAGATATGTCAGATC 123
QY 22 CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
Db 124 TGTGGTGATGATGTTGGACTCGCTGGAACCTGGAGATGCTTTGTCGCGGTGAATGAATGT 183
QY 42 ArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGluAlaCys 61
Db 184 GCCTTCCCTGTGTCGCGCTTGTATGATGATGACAGAGAAAGATGGAACCTCAGTGTTC 243
QY 62 LeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlu 81
Db 244 CCTCAATGCAAGACTAGATTTCAGACACACAGGGGGAGTCTCTGTTGAAGAGATGAA 303
QY 82 GlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGlu 101
Db 304 GATGAGATGATGTTGATGATATCGAATGAGTTCATTAATACGCCAGGGAGCTAACAG 363
QY 102 AspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySer 121
Db 364 GCGAGACACCAACGCCCATGGCGAAGAGTTCCTCTCT----- 399
QY 122 GlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp--- 140
Db 400 -----TCCTCTAGACATGAATCTCAACCAATTCCTCTCTCACCATTGGCCAT 447
QY 141 -----SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSer 158
Db 448 ACGGTTTCTGGAGAGATTCCGACGCCCTGATACACAATCTCTGCGAACTACA-----TCA 501
QY 159 GlyGluIleProGlyAlaSerProAspHisMetMetSerProThrGlyAsnIleSer 179
Db 502 GGTCTCTTGT--GGTCTCTCTGACAGGAATGCTATTTCATCTCCA-----TATATTGAT 552

```

QY 179 ArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe---Ser 197
DB 553 CCACGGCAACCTGTCCTGTA-----AGAACTCGTGACCCGTCAAAGACTTTGAACCTCT 606
QY 198 GlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAsp 217
DB 607 TATGGCTTGGTAATCTTGAACGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 218 LysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAla 237
DB 667 AAAAATATGTTACAGATGACTGGT-----AAATACCATCAAGGGAAAGAGAGA 714
QY 238 ThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThr 257
DB 715 ---GAANTGAA--GGGACTGGTTCCATGCGGAGAACTCCAATGGCTGATGATACA 768
QY 258 ArgGlnProLysSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMet 277
DB 769 GCTCTTCCTATGAGTCGTCGTGTCCTATCCCATCTCTCGCCTAACCCCTTATCGGGTT 828
QY 278 ValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnPro 297
DB 829 GTGATATATCTCCGGCTTATCATCTGTGTTCTTTCGAAATATCGTACAACTCACCCCT 888
QY 298 ValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeu 317
DB 889 GTGAAAAATGCATATCCTTTGGTTGACCTCGGTTATCTGTGAGATCTGGTTTGCATTT 948
QY 318 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTrpLeuAsp 337
DB 949 TCTTGCTCTTGTAGTCAGTTTCCAAATGGTACCCCATTTAAACAGGAGACTTATCTTGAC 1008
QY 338 ArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIle 357
DB 1009 COTCTCGCTTAAGATATGATCGACGCGTGAACCATCACAGCTCGTCTCTTGTATGG 1068
QY 358 PheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeu 377
DB 1069 TTTGTTAGTACAGTGGACCAATTAAGAGAGCCCTCCCTGTTTACAGCAACACACAGTTCTC 1128
QY 378 SerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGly 397
DB 1129 TCGATTCCTTCTGTGGACTACCCGCTAGATAAAGTAGCCCTGTTATGTTTCAGATGATGT 1188
QY 398 AlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpVal 417
DB 1189 TCAGTAGTCTTACCTTTGAATCCCTTCTGAAACCGCTGAGTTTGCAGAAAGAAATGGGTA 1248
QY 418 ProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLys 437
DB 1249 CCATTTTGCAGAAATTCACACATTCGAACCTAGGGCCCTGGAATTCATTTTGGCCAGAG 1308
QY 438 IleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLys 457
DB 1309 ATAGATTACTTTGAAGGCAAGATCAACCGCTCTTTTGTAAAGACGACGAGCTATGAAG 1368
QY 458 ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysVal 477
DB 1369 AGAGAGTATGAAGAGTTTAAGTAGAGATTAATGCTCTTGTTCGCAAGACAGAAATC 1428
QY 478 ProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAsp 497
DB 1429 CCTGAAGAAGGCTGACCAATGACGATGGTACTCCCTGGCCTGGTAAACAACACATAGAGAT 1488
QY 498 HisProGlyMetIleGluValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsn 517
DB 1489 CATCTGGAATGATACAGGTGTTCTTAGCCCATAGTGGGGTCTGGATACCGATGGAAT 1548
QY 518 GluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLys 537
DB 1549 GAGTCGCTAGACATCATCTATGTTCTCGTGAAGAGCGCCTGGATTTCAACACACACAAA 1608
QY 538 LysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyr 557

DB 1609 AAGGCTGGAGCTATGAATGCATTCGCTGATTCGCTGTTCTTACCAATGGAGCATAT 1668
QY 558 MetLeuAsnLeuAspCysAspHisTyrIleLeuAsnSerLysAlaValArgGluAlaMet 577
DB 1669 CTTTGAACGTCGATGTCATTCATTAACAGTAAGGCTATTAAAGAAGCTATG 1728
QY 578 CysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArg 597
DB 1729 TGTTCATGATGAGCCCGCTATTGGAAGAAGTCTGCTATGTCCAGTTCCTCAACGT 1788
QY 598 PheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIle 617
DB 1789 TTTGACGCTATGATTCACGATCGATATGCCACAGAAATATAGTCTTTTCGATAT 1848
QY 618 AsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPhe 637
DB 1849 AACATGAAGGGTTCGATGTCATCCAGGGTCCAGTATATGTGGTACTGGTTGTGTTT 1908
QY 638 AsnArgThrAlaIleTyrGlyTyrGluProIleLysAlaLys-----LysProGly 655
DB 1909 AATAGCAGGCTCTATATGGGTATGATCTGTTTGACGGAAGAAGATTTAGAACCAAT 1968
QY 656 PheLeuAla---SerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSer 674
DB 1969 ATATTGTCAGAGAGCTGTTGGGGTCAAGGAAGAAAGGTAAAGGTAGCAAGAAGTATAAC 2028
QY 675 SerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGlu 694
DB 2029 TACGAAAAGAGGAGGAGCATCAACAGAGTGAATCTCAATGTCTCCACTTTTCAATATGGAG 2088
QY 695 AspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSer 714
DB 2089 GACATCGATGAGGGTTTGA-----GGTTATGATGAGAGGTCTATTCTAATGTC 2142
QY 715 GlnMetSerLeuGluLysArgPheGlyGlnSerAlaIlePheValAlaSerThrLeuMet 734
DB 2143 CAGAGGAGTGTAGAGAAGCGTTTGGTCAGTCGCGGTATTATTTCGCGCAACCTTCATG 2202
QY 735 GluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHis 754
DB 2203 GAACAAGCGCGCATCTCCACCAACCACTCCGCTACTCTTCTGAAGAGGCTATTCTAT 2262
QY 755 ValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyr 774
DB 2263 GTTATAAGCTGTGTTACAGACAGACTGAATGGGGCAAGAGATTTGGTTGATCTAT 2322
QY 775 GlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSer 794
DB 2323 GGTTCGTCGTCGGAAGATATTCTTACTGGGTTCAAGATGTCATGCCCGGGTTGGATATCG 2382
QY 795 ValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAsp 814
DB 2383 ATCTACTCAATCTCTCCACGCCCTCGCTCAAGGGATCTGCACCAATCAATCTTTCTGAT 2442
QY 815 ArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHis 834
DB 2443 CGTTTGAACCAAGTCTTCGATGGCTTTGGATCTATCGAGATTTCTTTAGCAGACAT 2502
QY 835 CysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
DB 2503 TGTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2562
QY 855 AsnThrThrIleTyrProLeuProLeuLeuValTyrCysIleLeuProAla 874
DB 2563 AACACATCGTCCTATCTTATACATCCATCCCTCTTATTGGGTATTGATTTCTTCCCGCT 2622
QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrp 894
DB 2623 TTTTGTCTCATCCGACAGATTCATCATCCCGAGATTAAGCAACTACCGCAGATTTGG 2682
QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 914

Db 2683 TTCATTCTACTTTCATCTCAATGCTGTGTGACTGGAATCTCTGGAGCTGAGATGGAGCGGT 2742
 Qy 915 ValGlyIleAspGluThrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAla 934
 Db 2743 GTGAGCATGTAGGATGTGGTGGAGAACGAGCAGTCTCTGGGTCAATGTGGTGGACATCCGCC 2802
 Qy 935 HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
 Db 2803 CATCTTTTGTCTCTCCAGGTCTACTTAAGTCTCTGTGTGTATGACACCACTTC 2862
 Qy 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuThrMetPheLys 974
 Db 2863 ACCGTTTACATCTAAAGCCACAGACGAGATGGGATTTTGCAGAACTCTACATCTCAAA 2922
 Qy 975 TrpThrThrLeuLeuLeuProProThrThrThrIleLeuLeuIleAsnMetValGlyValVal 994
 Db 2923 TGGACAGCTCTCTCAITCCACCAACACCGCTCTACTTGTGAACCTCATAGGATGTG 2982
 Qy 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLys 1014
 Db 2983 GCTGGTGTCTTTAAGCTGTAAACAGTGGCTTACCAGTCTGGGTCCGCTTTTCGGGAAG 3042
 Qy 1015 LeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
 Db 3043 CTCCTCTGGCTTATGGGTTATTGCCCATCTCTACCCCTTTCTTGAAAGTCTGTGGGA 3102
 Qy 1035 ArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePhe 1054
 Db 3103 AGACAAACCGAACCAACCATCGTCATTGTCTGGTCTGTCTCTCGCCTCCATCTTC 3162
 Qy 1055 SerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsn 1072
 Db 3163 TCGTTGCTTGGGTGAGATCAATCCCTTTGTGGAC-----GCCAATCCCAAT 3210

RESULT 3

US-09-221-013A-11
 ; Sequence 11, Application US/09221013A
 ; Patent No. 6495740
 ; GENERAL INFORMATION:
 ; APPLICANT: Arioli, Antonio
 ; APPLICANT: Williamson, Richard E.
 ; APPLICANT: Betzner, Andreas S.
 ; APPLICANT: Peng, Liangcai
 ; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
 ; FILE REFERENCE: 96-98
 ; CURRENT APPLICATION NUMBER: US/09/221,013A
 ; CURRENT FILING DATE: 1998-12-23
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00402
 ; PRIOR FILING DATE: 1997-06-24
 ; PRIOR APPLICATION NUMBER: AU P00699
 ; PRIOR FILING DATE: 1996-06-27
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 11
 ; LENGTH: 3673
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (71)..(3313)
 US-09-221-013A-11

Alignment Scores:

Pred. No.: 0 Length: 3673
 Score: 3959.50 Matches: 732
 Percent Similarity: 80.71% Conservative: 138
 Best Local Similarity: 67.90% Mismatches: 169
 Query Match: 68.53% Indels: 39
 DB: 4 Gaps: 15

US-09-900-237A-30 (1-1080) x US-09-221-013A-11 (1-3673)

Qy 2 AspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIle 21

Db 143 GATGGCGGACCAACCTTTGAAGAATATGAAT-----GGCCAGATATGTCAGATC 193
 Qy 22 CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
 Db 194 TGTGGTGATGATGTGGACTCGCTGAAATGAGATGCTCTTTGTCGGCTGTAATGAATGT 253
 Qy 42 ArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCys 61
 Db 254 GCCTTCCTCGTGTCTGGCCCTTGTATGAGTACGAGAGAAAGATGAACACTCAGTGTTC 313
 Qy 62 LeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlu 81
 Db 314 CCTCAATGCAAGACTAGATTTCAGACGACACACAGGGGAGTCTCTCGTGTGAAGAGATGA 373
 Qy 82 GlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGlu 101
 Db 374 GATGAGGATGATGTGTGATGATATCGAAGATGAGTTCAATTACGCCAGGAGCTTAACAAG 433
 Qy 102 AspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySer 121
 Db 434 GCGAGACCAACGCCCATCGCGAAGAGTGTTCCT----- 469
 Qy 122 GlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp--- 140
 Db 470 -----TCCTCTAGACATGAATCTCAACCAATTCCTCTCTCACCATGGCCAT 517
 Qy 141 -----SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSer 158
 Db 518 ACGGTTTCTGGAGAGATTCGCACGCTGATACACAACTCTGCGAAGTACTACA-----TCA 571
 Qy 159 GlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSer 178
 Db 572 GGTCTCTTGTG---GGTCTCTCTGACAGGAATGCTATTTCATCTCCA-----TATATTGAT 622
 Qy 179 ArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe---Ser 197
 Db 623 CCACGGCAACCTGTCCCTGTA-----AGATCTGTGGACCGCTCAAAAGACTTGAACCTCT 676
 Qy 198 GlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAsp 217
 Db 677 TATGGGCTTGGTAATGTGTGACTGGAAGAAAGAGCTTGAAGGCTGGAAGCTGAAGCAGGAG 736
 Qy 218 LysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAla 237
 Db 737 AAAAATATGTTACAGATGACTGGT-----AAATACCAGTGAAGGAAGAGGAGA 784
 Qy 238 ThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThr 257
 Db 785 ---GAATTTGNA---GGGACTGGTTCCTCAATGGCGAAGNACTCCAAATGGCTGATATACA 838
 Qy 258 ArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMet 277
 Db 839 CGTCTCTATGAGTCGTGTGTGCTTATCCATCTCTCGCCCTTAAACCCCTTATCGGGTT 898
 Qy 278 ValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnPro 297
 Db 899 GTGATTATTCTCGGCTTATCATCTCTGTGTTTCTTGTCAATATCGTACAACTCACCCCT 958
 Qy 298 ValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeu 317
 Db 959 GTGAAAAATGCATATCTTTGTGGTTCACCTCGGTTATCTGTGAGATCTGGTTTGCATTT 1018
 Qy 318 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAsp 337
 Db 1019 TCTTGGCTTCTTGATCAGTTTCCCAATGGTACCCTTAACAGGAGACTTATCTTTGAC 1078
 Qy 338 ArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIle 357
 Db 1079 CGTCTCGCTATAAGATATGATCGAGACGGTGAACCATCACAGCTCGTTCCTGTGTGATGTG 1138
 Qy 358 PheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeu 377

Db 1139 TTTGTTAGTACAGTGGACCCATTGAAAGAGCCCTCCCTTGTGTACAGCAAAACACAGTTCTC 1198
QY 378 SerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGly 397
Db 1199 TCGATTCTTTCTGTGAGCATACCCGGTAGATAAAGTAGCCCTGTATGTATTCAGATAGGT 1258
QY 398 AlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrVal 417
Db 1259 TCAGCTAGCTTACCTTTGAATCCCTTCTGAAACCGCTAGCTTCCAAAGAAATGGGTA 1318
QY 418 ProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLys 437
Db 1319 CCATTTTGCAGAAATCAACATTCGAACCTAGGGCCCTGAAATCTTATTTTCCCAAGAG 1378
QY 438 IleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLys 457
Db 1379 ATAGATTACTTGAAGACCAAGATCCAACCGCTTTTGTGTAAAGAGCGAGCTATGAAG 1438
QY 458 ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysVal 477
Db 1439 ACAGAGTATGAAGAGTTTAAAGTAGGATAAATGCTCTTGTGTCACAAAGCACAGAAATC 1498
QY 478 ProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGluAsnAsnThrArgAsp 497
Db 1499 CCTGAAGAGGCTGGACATGAGATGGTACTCCCTGGCCCTGGTAACACACTAGAGAT 1558
QY 498 HisProGlyMetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsn 517
Db 1559 CATCTCGAATGATACAGGTGTTCTTAGGCCATAGTGGGGTCTGGATACCGATGGAAAT 1618
QY 518 GluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLys 537
Db 1619 GAGCTGCCTAGACTCATCTATCTTCTGCGTGAAGAGCGCTGGATTTCACACACCACAAA 1678
QY 538 LysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyr 557
Db 1679 AAGGCTGGAGCTATGAATGCATTCGATCGTGTATCTGTTGTTCTTACCAATGGAGCATAT 1738
QY 558 MetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMet 577
Db 1739 CTTTGTGAAGTGGATTGTGATCATTTACTTTAATACAGTAAAGCTATTAAAGAGCTATG 1798
QY 578 CysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArg 597
Db 1799 TGTTTTCATGATGCACCGCTATTGGAAAGAGTGTCTATGTCTCAGTTCCCTCAACGT 1858
QY 598 PheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIle 617
Db 1859 TTTGACGGTATTGATTTCACCATCGATATGCCAAGAGATATAGTCTTTTTTCGATATT 1918
QY 618 AsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPhe 637
Db 1919 AACATGAAGGGTTGGATGGTATCCAGGTCAGTATATGTGGTACTGTGTTGTGTTT 1978
QY 638 AsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----LysProGly 655
Db 1979 AATAGCAGGCTCTATATGGGTATGATCTCTGTTTTCAGCGAAGAGATTTCAGCAACAAAT 2038
QY 656 PheLeuAla---SerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSer 674
Db 2039 ATATTGTCAAGAGCTGTTCGGGTCAGGAAGAAGGTAAAGTAGCAAGAGATATAC 2098
QY 675 SerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGlu 694
Db 2099 TACGAAGAAGAGGAGGCATCAACAGAGTGACTCCAATGCTCCACTTTTCAATATGGAG 2158
QY 695 AspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSer 714
Db 2159 GACATCATGATGAGGGTTTTCGAA-----GGTTATGATGATGAGAGGTCTATTCTAATGTCC 2212
QY 715 GlnMetSerLeuLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMet 734
Db 2213 CAGAGGAGTGTAGAGAGGTTTTGGTCACTGCCGGTATTATTATGCGGCAACCTTCATG 2272

QY 735 GluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHis 754
Db 2273 GAAACAGGCGGATTCACCAACCAACCAATCCCGTACTCTTCTTGAAGAGGCTATTCTAT 2332
QY 755 ValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyr 774
Db 2333 GTTATAAGCTGTGGTTACGAACAAAGACTGAATGGGCAAGAGATTGCTTGGATCTAT 2392
QY 775 GlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrPargSer 794
Db 2393 GGTTCGGTACGGAAGATATCTTACTGGGTTCAAGATGCATGCCCGGGTTGGATATCG 2452
QY 795 ValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAsp 814
Db 2453 ATCTACTGCAATCCCTCCAGCCCTGGTTCAGGGATCTGCACCAATCAATCTTCTGAT 2512
QY 815 ArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHis 834
Db 2513 CGTTTTGAACCAAGTTCTTCGATGGGCTTTGGGATCTATCGAGATTCTTCTTAGCAGACAT 2572
QY 835 CysProLeuTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
Db 2573 TGTCTCTATCTGGTATGGTTACCATGGAAGTTGAGACTTTTGGAGAGGATCGCTTATATC 2632
QY 855 AsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
Db 2633 AACACCATCGTCTATCTTATCATCTCCATCCCTCTTATTCGCTATGTATTTCTCCGCT 2692
QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyr 894
Db 2693 TTTTGTCTCATCCGACAGATTCTATACCCGAGATAAGCAACTACGCGAGTATTTGG 2752
QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 914
Db 2753 TTCTATCTACTCTTCTCATCTCAATTGCTGTGACTTGGAAATCCTGGAGCTGAGATGGAGCGT 2812
QY 915 ValGlyIleAspGluTyrTrpArgAsnGluGlnPheTyrValIleGlyIleSerAla 934
Db 2813 GTGACATTTGAGATTGGTGGAGGAACGAGCAGATTCTGGTCAATTTGGTGACATCCGCC 2872
QY 935 HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
Db 2873 CATCTTTTGTCTGTCTTCCAGGTTCTTAAAGTTCTTGTCTGTGTATCGACACCACTTC 2932
QY 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
Db 2933 ACCGTTTACATCTAAAGCCACAGCAAGATGGGATTTTGCAGAACTCTACATCTTCAA 2992
QY 975 TrpThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyValVal 994
Db 2993 TGGACAGCTCTTCTCATTTCCACCAACCCGCTTACTTGTGAACCTCATAGCATTTGTG 3052
QY 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLys 1014
Db 3053 GCTGTGTCTCTTATGCTGTAAACAGTGGCTACCAAGTCCGTGGGGTCCGCTTTTCGGAAG 3112
QY 1015 LeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
Db 3113 CTCCTCTTCGCTTATGGGTATTGCCCATCTTACCCCTTTCTTGAAGGCTGTGTGGGA 3172
QY 1035 ArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePhe 1054
Db 3173 AGACAAAACCGAAACCAACCAATCGTCTTGTGTGTCTGTCTTCTCGCTCCATCTTTC 3232
QY 1055 SerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsn 1072
Db 3233 TCGTTGCTTGGGTGAGGATCAATCCCTTTGTGGAC-----GCCAATCCCAAT 3280

RESULT 4

US-09-221-013A-7

; Sequence 7, Application US/09221013A

; Patent No. 6495740


```
QY 547 ArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyr 566
Db 1880 CGAGTCTCTGCTGTTCTATCAACCGCTCCTTACCTCTTTAATGTCGATGTCATCACTAC 1939
QY 567 IleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGly 586
Db 1940 ATCAACAACAGCAAGCAATAGAGAAATCTATGTGTTTCATGATGGACCCGCAATCGGGA 1999
QY 587 ProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArg 606
Db 2000 AAGAAGTTTGTATGTTTCAGTTCGCGAGAGATTGATGGATTGATAGATAGATAGA 2059
QY 607 TyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGln 626
Db 2060 TACTCAACCGTAACGTTGTGTTCTTGATATTAAATGAAAGGCTTGTATGGGATACNA 2119
QY 627 GlyProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGlu 646
Db 2120 GGACCGATATATGTCGGACAGGTTGTGTGTTTAGAAAAACAGGCTCTTTATGTTTGAT 2179
QY 647 ProProIleLysAlaLysLysProGlyPheLeuAlaSer----- 659
Db 2180 GCACCAAGAGAGAAACCCACAGCAAAACCTGTAACTGTGGCTTAAATGGTGTGT 2239
QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2240 TTGTGTTGGGTTGAGAAAG-----AAGAGTAAACGAAAGCCACAGATAAGAAACT 2293
QY 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
Db 2294 AACACTAAA-----GAGACTTCAAGACAGATTTCATGCGCTAGAGAAATGTCACGAAGGT 2347
QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2348 GTTATCGTCCCGAGTGTCAAAATGTTGAGAAAGAGATCTGAAGCAACACAAATTTGAG 2407
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2408 AAGAAGTTTGGACAATCTCCGGTTTTCGTTGCTCTGCTGTCTACAGAACGGTGGAGTT 2467
QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2468 CCCCCTAAGCAAGCCCGCATGTTGTTTAAAGAGAGCCATTCAAAGTTATTAGCTGCGGG 2527
QY 760 TyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db 2528 TACCAGATAAACCAATGGGAAAGAGATCGGGTGATTTATGGATCGGTGACTGAA 2587
QY 780 AspileLeuThrGlyPheLysMethisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2588 GATATCCTGACGGTTTCAAGATGCATTCGCCATGGATGGAGATCTGTGACTGATGCT 2647
QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2648 AAGCGTGAGCGCTTTTAAAGGATCTGCTCTATTAACTTGTACAGTCTGCTTCATCAAGTT 2707
QY 820 LeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyr 839
Db 2708 CTACGTTGGGCTCTGGCTCTGTAGATTTCTTGACGAGACATTCGCGATATGGTAT 2767
QY 840 GlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2768 GGTATGTTGGTGGTGTAAATGGTTGGAGAGATTCTCTTACATCAACTCTGCTCTAT 2827
QY 860 ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2828 CCTCGACTTCACTTCCATGATGCTGCTATTGCTCTCCCGCGGTGTTGTTACTCACA 2887
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe 899
Db 2888 GGAAATTCATGCTCCCTGAGATAGCAACTACGAGGATATACTCTTCATGCTCATGTT 2947
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
```

```
Db 2948 ATATCATAGCAGTAACCTGGAATCCTCGAAATGCAATGGGAGGTGTCGGAATCGATGAT 3007
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 3008 TGGTGGAGAAACAGCAGCTTTTGGGTAAATCGGAGGGGCTCTCCGATCTATTGCTCTG 3067
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3068 TTTCAAGGTTTCTCAAAGTTCTAGCCGAGGTAAACAGAAATTTACAGTCACTTCAAAA 3127
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 3128 GCAGCAGAC---GATGGAGCTTCTCTGAGCTTTACATCTTCAAGTGGCAACTTTGTTG 3184
QY 980 IleProProThrThrIleLeuIleLeuMetValGlyValValAlaGlyThrSerTyr 999
Db 3185 ATTCTCTCGACACCACTTCTGATCATTAACTCATTTGGAGTTATTGTCGGCGTTTCTGAT 3244
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3245 GCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGAGACTTTCTTCGCTCTT 3304
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3305 TGGGTCAATTGTTTATACCAATCTCTCAAGGGAATGCTTGGGAAGCAACAAATG 3364
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3365 CCTACGATTTATGTCGTCTGTTCTATTCTTTAGCTTCGATCTTACACACTCTTTGGGTC 3424
QY 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3425 AGAATTAACCCGTTTGTGGCTAAA---GGGGACCAAGTCTTGGAGATCTGTGCTGAT 3481
QY 1080 Cys 1080
Db 3482 TGT 3484

RESULT 5
US-09-221-013A-3
; Sequence 3, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 8411
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-3

Alignment Scores:
Pred. No.: 0 Length: 8411
Score: 3518.50 Matches: 743
Percent Similarity: 55.50% Conservative: 140
Best Local Similarity: 46.70% Mismatches: 158
Query Match: 60.83% Indels: 553
DB: 4 Gaps: 28

US-09-900-237A-30 (1-1080) x US-09-221-013A-3 (1-8411)
```


QY 16 GlyAspValCysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPhe 35
Db 2928 GGCAGATGTCAGATCTGTGGTGGATGATGTCGCTGAACTGGAGATGCTTTT 2987
QY 36 ThrAlaCysAspValCysArgPheProValCysArgProCysTyrGluHisGluArgLys 55
Db 2988 GTCGGTGTAAATGAATGTCCTCCCTGTCGTGTCGGCTTGTCTATGATGACGAGAGAAA 3047
QY 56 GluGlyThrGlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArg- 72
Db 3048 GATGGAACCTCAGGTGTCCTCAATGCAAGACATAGATTCAGACGACACAGGGGTTCAGTTG 3107
QY 72 - 72
Db 3108 TCTTTTCTTTTGTGGCAATTGCTATATATATGGAATTTCTCTTTTGTCTTTGCTGT 3167
QY 73 - - - - -GlySerProAlaIleArgGlyGluGluGly 82
Db 3168 TGTGTTGAACAATTTTGGAAATTTTCCAGGAGT-CCTCGTGTGAGGAGATGAGAT 3226
QY 83 AspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAsp 102
Db 3227 GAGGATGATGTCATGATATCGAATGATGAGTTCATTAACGCCAGGAGGCTAACAGGCG 3286
QY 103 GlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySerGly 122
Db 3287 AGACACCAACGCCATGCGGAGAGTTCCTTCCT- - - - - 3319
QY 123 AsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyr- 139
Db 3320 - - - - -TCCCTAGACATGAATCTCAACCAATTCCTCTTCACCCCATGGCCATACG 3370
QY 140 - - - - -AspSerGlyGlu- - - - -IlePro 145
Db 3371 GTAGGACCTACATTTTCCCTTTAGACTCTAGATGATTTGTATTACTCAATAATCCCT 3430
QY 146 Arg- - - - - 146
Db 3431 AGAGTGGTCATTTATTACTTACTTATTCACGTTAAATGTTATATGTGAACAAATCTTAACAG 3490
QY 147 - - - - -GlyTyr- - - - -Va 149
Db 3491 AATTTTCTGTAGTACATGTCATCCAAATTAAGAAATAATAAGATGTTGTAGT 3550
QY 149 lProSerValThrAsn- - - - -SerGlnMetSerGlyGluIlePr 162
Db 3551 TGTGTCGTGTTTCAATAGATTCATGACCTTTTCTATACACAGGTTTCTGGAGAGATT- 3608
QY 162 oGlyAlaSerProAspHisMetMet- - - - -SerProThrGlyAsnIleSe 178
Db 3609 - - - - -CGCACGCTGATACACAATCTGTGCGAATACATCAGTCCCTTTGGGTCTCTCTGA 3664
QY 178 rArgArgAla- - - - -ProPheProTyrValAsn-HisSer- - - - - 189
Db 3665 CAGAAATGCTATTTCATCTCCATATATATGATCCAGCGCAACCTCGTATTCCATATGTTTT 3724
QY 189 - - - - - 189
Db 3725 CCCTTGTGCAGTGGTCTTTGTTAAATGTGATTCCTATTCTTTTACACATATATATT 3784
QY 190 - - - - -ProAsn 192
Db 3785 TTGTGTACCGTAATGATAGTCCCGCTAAATAATGTCAGTCCCTGTGAATCTGGGACC 3844
QY 192 roSerArgGluPhe- - - - -SerGlySerIleGlyAsnValAlaTlPylsGluArgValAsp 211
Db 3845 CGTCAAAAGACTTGAACCTCTTATGGGCTTGTATATGTTGATCGAAGAAAGAGTTGAAG 3904
QY 211 lyTrpLysMetLysGlnAspLysGlyAlaIleProMetThr- - - - - 224
Db 3905 GCTGGAAGCTGAAGCAGGAGAGAAAATATGTTACAGATGACTGGTAAATACCATGAAGGA 3964

QY 225 - - - - -AsnGlyThrSerIle- - - - - 229
Db 3965 AAGGAGAGAAAATTGAAGGACTGGTTCATGCGAAGAACTCCAAATTAAGTGAAA 4024
QY 230 - - - - -AlaProSerGlu- - - - -GlyArgAlaAlaThrAspI 240
Db 4025 TACTAGACCAATATCTTTATTGTCCAACCTCAACAGCTCTTGGCCGTGATGCTAATAAC 4084
QY 240 leAspAlaSerThrGluTyrAsnMetGluAspAlaLeu- - - - - 252
Db 4085 ACTCTTG-CTTCTTATTATGATGATAGACATAAATTAAGTATCTGCTTTGTTACATTT 4143
QY 253 - - - - -LeuAsn- - - - -AspGluThrArgGlnProL 261
Db 4144 GTTTCCTTCACATCAATATGTTCTCGTACTTACAGGCGTGATGATACGCTTCTCTA 4203
QY 261 euSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 281
Db 4204 TGAGTCGTGGTGGCTATCCCATCTTCTCGCTAACCCCTTATCGGGTTGTGATTATTC 4263
QY 281 euArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnA 301
Db 4264 TCCGGCTTATCATCTTGTGTTCTTCTGCAATATCGTACACTCACCCCTGTGAAAATG 4323
QY 301 laTyrProLeuTlPLeuLeuSerValIleCysGluIleTlPLeuAlaLeuSerTlPLeu 321
Db 4324 CATATCCTTTGTGGTGGCTCGGTTATCTGTGAGATCTGGTTGTCATTTCTTGGCTTC 4383
QY 321 euAspGlnPheProLysTlPLeuPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 341
Db 4384 TTGATCAGTTTCCCAATGTTAGCCCATTAACAGGAGACTTATCTTGAACGCTCGCTA 4443
QY 341 eu- - - - - 341
Db 4444 T-AAGGTTGGTCTTTAAAGTTTATACATCCCTACTCTCATCTCTCTTTTATGTTAACT 4502
QY 341 - - - - - 341
Db 4503 TGATATCTTCTATCACAGTTTTCGATAGTTGACTTTTCCCCCTGTAAAATTAATTTAAA 4562
QY 341 - - - - - 341
Db 4563 TTTAGACAATGGTGCATCTGAATTTTGATTATGATATATCTTAAGAAGATTATGATTGTA 4622
QY 341 - - - - - 341
Db 4623 AATCTTGAAATTTAGTAGAAGAACCATCTGCAATCTACTGACCATGTGAAGTTCCGACTA 4682
QY 341 - - - - - 341
Db 4683 GACTATGATAGAAGCATGCCAAGTGGAGTGTATTATTAAGATAGAGCTTAGCTATTATACT 4742
QY 342 - - - - -ArgTyrAspArgGluGluP 349
Db 4743 GATTTTATATGTGTTTGAATTTTGTGTTTCTTATTGTAGATATGATCGAGACGGTGAAC 4802
QY 349 roSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProp 369
Db 4803 CATCACAGCTCGTTCCTCGTTGATGTGTTGTTGTAGTACAGTGGAGCCCAATTGAAAGAGCCTC 4862
QY 369 rolleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysV 389
Db 4863 CCCTTGTTCAGCAAAACACAGATTCCTCGATTCTTCTGTGGACTACCCGGTAGATAAAG 4922
QY 389 alSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluT 409
Db 4923 TAGCCTGTATGTTTTCAGATGATGTTTCAGCTATGCTTACCTTTGAATCCCTTTCTGAAA 4982
QY 409 hrSerGluPheAlaArgLysTlPValProPheValLysLysTyrAspIleGluProArgA 429
Db 4983 CCGCTGAGTTTGCAAAAGAAATGGGTACCAATTTTGCAAGAAATTCACATTTGAACCTAGGG 5042
QY 429 laProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSerP 449

Db 5043 CCCCCTGAATCTTATTTGCCCAGAGATAGATTACTTGAAGCACAAGATCCACCGCTCTT 5102
QY 449 heValLysAspArgAlaMetLys 457
Db 5103 TTGTTAAAGAGCGAGCTATGAAGGTCAITTTGAAAAGTCCACCTCTCTCATCCATA 5162
QY 458 459 -ArgGluTyrGluGluP 463
Db 5163 CGGCAAGAGATTGACTGACTTTTTCTTTGGTTTGTATTGACAGAGAGATGAAGAGT 5222
QY 463 heLysileArgileAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTyrPI 483
Db 5223 TTAAGAGTGAAGATAATGCTCTTGTGCCAAGCAGAGAAATCCCTGAAGAGGCTGA 5282
QY 483 leMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIle- 502
Db 5283 CAATGAGGATGGTACTCCCTGGCCTGGTAAACAACACTAGAGATCATCTCGGAATGATAC 5342
QY 502 503 504
Db 5343 AGGTACAGTGGCAATCCCTTGATTGTGACAGAGGATACGTAAGGAACAATGTTT 5402
QY 503 504 -GlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluG 516
Db 5403 ACATCGTTTTGTTTCAATTTCAAGGTGTTCTTAGGCCATAGTGGGGTCTGGATACCGATG 5462
QY 516 lysAsnLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisH 536
Db 5463 GAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAAGCGGCTCGAATTTCAACACC 5522
QY 536 isLysLysAlaGlyAlaMetAsnAlaLeuVal- 546
Db 5523 ACAAAAAGCTGGAGCTATGATGATGCTTGTGTTTAACTTTCAGAACTCTATGTGTGCC 5582
QY 546 547 548 549 550
Db 5583 TCTATTTTATCTCTGTTTCACTGCTTAAGAAACGTTCTTCTGTGTAGCCGCTTCTCA 5642
QY 546 547 548 549 550
Db 5643 CATTCCTTTTTTTAGGCTARGTGTCTCTCTAATTTAGTATCTTTTACTTTTGACAG 5702
QY 547 548 -ArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHis 565
Db 5703 ATCCGTGTATCTGCTGTTCTTACCATATGAGATATCTTTTGAACGTGGATGTGATCAT 5762
QY 566 TyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeu 585
Db 5763 TACTTTTAATAACAGTAAGGCTATTAAAGAAAGCTATGTGTTTCATGATGACCCCGCTATT 5822
QY 586 GlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyTleAspArgAsnAsp 605
Db 5823 GGAAGAAGTCTGCTATGTCAGGTCCCTCAACGTTTTTGACGGTATTGATTTGACAGT 5882
QY 606 ArgTyrAlaAsnArgAsnThrValPhePheAsp- 616
Db 5883 CGATATGCCAACAGGAATATAGTCTTTTTCGATGTGAGTATCACATTCGCCCATTTGCTTTT 5942
QY 616 617 618 619 620
Db 5943 GTTTCCTTTTTTTCATATTTTGGTGGATTACTCGTTTCTGCTATGGCCTGACTTGA 6002
QY 617 618 -IleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
Db 6003 TATTTGTTCTCTTGGGCAGATTAAACATGAAGGGTTGGATGTATCCAGGGTCCAGTATA 6062
QY 630 rValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleIly 650
Db 6063 TGTGGGTACTGTTGTTGTTTAAATAGCAGGCTCTATATGGGTATGATCCTGTTTGAC 6122
QY 650 sAlaLys-----LysProGlyPheLeuAla---SerLeuCysGlyGlyLysLysAl 667
QY 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

Db 6123 GGAAGAAGATTTTAGAACCAATATTTATTGTCAGAGCTGTTGCGGGTCAAGGAAGAAAGG 6182
QY 667 aSerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSe 687
Db 6183 TAAAGTAGTACGAAGAAGTATTAACAGAAAGAGAGAGGCATCAACAGAAAGTGACTCCAA 6242
QY 687 rValProValPheAsnLeuGluAspIleGluGlyValGluGly- 702
Db 6243 TGCTCCACTTTTCAATATATGAGGACATCATGAGGGTTTTGAAGGTTTTGATTGAGCTGAT 6302
QY 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```
QY 945 ysValleuAlaGlylleAspThrAsnPhetThrValThrSerLysAlaAsnAspGluGlu 965
Db 7263 AGGTTCCTGCTGGTATCGACCAACATTCACCGTTACATCTAAAGCCACAGCAAGATG 7322
QY 965 lyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuLeuProProThrThrI 985
Db 7323 GGGATTTCGAGAACTCATCTTCAATGGACAGCTCTCTCATTCACCAACACCG 7382
QY 985 leLeuIlelleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyT 1005
Db 7383 TCCTACTGTGTGAACCTCATAGGCATTCGTGGTGTCTCTTATGCTGTAACACAGTGGCT 7442
QY 1005 YrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisL 1025
Db 7443 ACCAGTCGTGGGTCCTTCCTTCGGGAGCTCTCTTCGGCTTATGGGTATTCGCCATC 7502
QY 1025 euTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleV 1045
Db 7503 TCTACCTTCTTGAAGGCTCTGTGGAGACAAAACCGNACACCACTCGTCAATG 7562
QY 1045 alTrpAlaValLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProPheT 1065
Db 7563 TCTGGTCTGTCTTCGCTCCATCTCTCTGTTGCTTGGTTCAGGATCAATCCCTTTG 7622
QY 1065 hrThrArgLeuAlaGlyProAsn 1072
Db 7623 TGGAC-----GCCATCCCAAT 7639

RESULT 6
US-08-960-048-1
; Sequence 1, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIORITY FILING DATE: 1997-10-29
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-1

Alignment Scores:
Pred. No.: 0 Length: 3328
Score: 3477.00 Matches: 657
Percent Similarity: 73.60% Conservative: 129
Best Local Similarity: 61.52% Mismatches: 176
Query Match: 60.18% Indels: 106
DB: 3 Gaps: 15

US-09-900-237A-30 (1-1080) x US-08-960-048-1 (1-3328)
QY 18 ValCysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAla 37
Db 200 GTTTCGCACACTTCGTGTGACATGTTGGTGTGAATGTTAATGTGTGAACCTTTTGTGGCT 259
QY 38 CysAspValCysArgPheProValCysArgProCysTrpGluHisGluArgLysGluGly 57
Db 260 TGCATGAATGTAATTTCCCTATTGTGAAGAGTGTGTTTGAAGTATGATCTTAAGGAAGGA 319
QY 58 ThrGlnAlaCysLeuGlnCysLysThrLysTrpLysArgHisArgGlySerProAlaIle 77
Db 320 CGAAAAGCTTGTGCTGTGTGT-----GGTAGTCCATAT--- 352
```

```
QY 78 ArgGlyGluGluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAla 97
Db 353 -----GATGAAACCTTGTGGACGAT-----GTGAG 379
QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
Db 380 AAGCCACCGCGCATCATCGACATCGCTGCA-----CAT 415
QY 118 ThrGlySerGlyAsnValGly---HisProLysTyrAspSerGlyGluIleGlyLeu 136
Db 416 TTGAACAAGTCTCAGGATGTTGAATTCATGCAAGACATATCAGC-----AGTGTG 466
QY 137 SerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln 156
Db 467 TCTACATGATGATG-----481
QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisMetMetSerProThrGlyAsn 176
Db 481 -----481
QY 177 IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe 196
Db 482 -----GAAATG 487
QY 197 SerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGln 216
Db 488 GCTGAAGACAAATGGAAATTCGATTTGGAAGAACAGAGGTGGAAGTGGAAAGAAAGAAAG 547
QY 217 AsLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAla 236
Db 548 AACAGNAGAAGAGCTCGAACACT-----AAGTT 580
QY 237 AlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGlu 256
Db 581 GAAAGAGAGGTGAAATCCCACTGAGCAACAATAAGAGATAAAGCGACCGATGCT 640
QY 257 ThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArg 276
Db 641 TCC---CAGCCCTCTCGACTATAATCCAAATCCGAAAGACGACTTCACCAATCCGA 697
QY 277 MetValIleValLeuArgLeuValLeuSerIlePheLeuHisTyrArgLeuThrAsn 296
Db 698 ACCGTGATCATATGCGATGATGATCTTGGTCTTTCTTCCATTCGAGTAACAAAC 757
QY 297 ProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAla 316
Db 758 CCGCTTCACAGTGCTTTTGGACTGTGGCTCACTTCAGTCATATGTGAAATCTGGTTTGA 817
QY 317 LeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeu 336
Db 818 TTTTCCTGGGTGTGGATCAGTTCCTTAAGTGTATCTCTGTTAACAGGAAACATAGATT 877
QY 337 AspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAsp 356
Db 878 GACGACTATCTGCAAGATATGAAGAGAGAGTGAACCTGATGAACCTTGTGCAAGTTGAC 937
QY 357 IlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrVal 376
Db 938 TTCTTCGTGAGTACAGTGGATCCATTGAAAGAGCGCTCCATTGATTACTGCCAATACTGTG 997
QY 377 LeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAsp 396
Db 998 CTTTCCATCTCTGCTTGGACTCCCGGTGGATGAAGGTCTCTTGTATATATCTGATGAT 1057
QY 397 GlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrp 416
Db 1058 GTTTCGCGCCATGCTGACATTTGATCTCTAGTAGAAGACAGCCGACCTTTCAGAAAGTGG 1117
QY 417 ValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGln 436
Db 1118 GTTCCATCTCTGCAAAAAAATTTCCATGAAACCCCGGCGACCTGAGTTTTACTTCTCAG 1177
```

437 LysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMet 456
 1178 AAGATTGATTACTTGAAGATATAAGTGGAGCCCTCTTTGTAAAGACGTAGACTATG 1237
 457 LysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLys 476
 1238 AAAAGAGATTATGAAGAGATACAAAATTCGAATCAATGCTTTAGTTGCAAGAGGCTCAGAAA 1297
 477 ValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnThrArg 496
 1298 ACACCTGATGAAGAGTGAACATGAATGAACTCTTGGCCAGGAAATACCCCGGT 1357
 497 AspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly 516
 1358 GATCACCTGGCATGATTCAGGTTTCTTGGATATAGTGGTCTCGTGACATCGAAGA 1417
 517 AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHis 536
 1418 AATGAACCTTCCTCGACTGCTTTACGCTCTCTAGAGAGAAGACCTGGCTACCAACACCAC 1477
 537 LysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGln 556
 1478 AAAAGGCTGGGCTGAAATGCTTTGGTTAGGGTGCTCGCAGTCTTACAAATGCTCCC 1537
 557 TyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAla 576
 1538 TTCATCCTCAATCTTGATTGTGACCACTATGTTAAACAATAGCAAGCAGTTAGGGAGCA 1597
 577 MetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGln 596
 1598 ATGTGCTTCTTGATGGACCCCAAGTGTGTCGAGATGATGCTATGTGCAGTTTCTCCTCAA 1657
 597 ArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp 616
 1658 AGATTGTAGGGCATAGATAGGAGTGCATGATGCAATAGCAACACAGTTTCTTTGAT 1717
 617 IleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVal 636
 1718 GTTAAACATGAAGGCTTGATGGAAATCCAAAGGCCAGTTTATGTGGGAACAGGTTGTGT 1777
 637 PheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPhe 656
 1778 TTCAATAGCAAGACACTTTATGGCTATGCTCCACCTTCAATGCCAAGTTTTCGCCAAGTCA 1837
 657 LeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAsp 676
 1838 TCCTCTCATCTGTCTGTTGTCGCCGGCAGAGAAACCTTAAGATCCATCAGAG 1897
 677 LysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIle 696
 1898 CTTTATAGGGATGCAAAACGGGAAGAACTTGATGCTGTCATCTTTAACTTAGGGAAATT 1957
 697 GluGluGlyValGluGlyAlaGlyPheAspAsp---GluLysSerValLeuMetSerGln 715
 1958 GAC-----AATTATGATGATGATGAAGATCAATGTTGATCTCTCAA 1999
 716 MetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaLaserThrLeuMetGlu 735
 2000 ACAAGCTTGAGAAAACCTTTGGCTATCTTCAGCTCTTCATGAACTACACTAATAGAG 2059
 736 TyrGlyGlyValProGlnSerSerThrProGluSerLeuLysGluAlaIleHisVal 755
 2060 AATGAGGAGTGGCTGAATCTGCCAACCCCTCCACACTAATCAAGGAAGCAATTCATGTC 2119
 756 IleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGly 775
 2120 ATCACTGTGGCTATGAAGAAGAACTGCATGGGGAAGACATTTGGATGGATATATGT 2179
 776 SerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrPargSerVal 795
 2180 TCAGTCACTGAGGATATCTTAACCGCTTCAAAATGCATCGCCGAGGATGGATCGATT 2239
 796 TyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 815

2240 TACTGCATGCCCTTAAGGCCAGCAATTAAGGATCTGCACCCCAATCTGTCTGATCGG 2299
 816 LeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCys 835
 2300 TTGCACCAAGGTTCTTCGATGGGCTCTTGAAATCTGTGTAAATTTTCTTAAGCAGGATGTC 2359
 836 ProLeuTyrGlyTyr---GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
 2360 CCTCATGGTATGGCTTTGGAGGTGGTCTTAATAATGGCTTCAAGAGACTAGCATATATA 2419
 855 AsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
 2420 AACACCATTTGCTATCCTTTCCATCCTTCCACTCATTTGCTATTGTTCACTACAGCA 2479
 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyr 894
 2480 ATCTGCTCTCTCACAGGAAATTTATCATCAACAGCTCTCAACCTCGCAAGTGTCTTC 2539
 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGly 914
 2540 TTTCTTGCCCTTTTCTTTCCATTATCGTACTGCTGTCTCGAGCTCGATGGAGTGGT 2599
 915 ValGlyIleAspGluTyrTyrArgAsnGluGlnPheTyrValIleGlyGlyIleSerAla 934
 2600 GTACGATTTGAGGACTTATGGCGTAACGAGCAGTTTGGGTCTCATCGTGGCGTTTCAGCC 2659
 935 HisLeuPheAlaValPheGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
 2660 CATCTCTTTGGCGTCTTCCAAAGTTTCTTAAGATGCTTGGGGCATTTGACACCAACTTT 2719
 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
 2720 ACTGTCACTGCCAAGCAGCTGAT---GATGCAGATTTTGGTGACCTCTACATTGTGAAA 2776
 975 TrpThrThrLeuLeuLeuProProThrThrIleLeuIleIleAsnMetValGlyValVal 994
 2777 TGGACTACACTTCTAATCCCTCCAAACAACACTCTCATCGTCAACATGGTTGGTCTGCTT 2836
 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLys 1014
 2837 GCGGATTTCTCGATGCCCTCAACAAAGGTACGAAGCTTGGGGACCACTCTTTGGCAAA 2896
 1015 LeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
 2897 GTGTCTTTCTCTCTGGGTCTCATCTCCATCTTTATCCATTCCTCAAGGCTCTTATGGGA 2956
 1035 ArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIlePhe 1054
 2957 GCGCAAAACAGGACACCAACCACTTGTCTCTTGGTCAGTGTGTGGTCTTCTGCTCTTC 3016
 1055 SerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle--- 1073
 3017 TCTCTGTGTGGTTCGGATCAACCCGTTTGTGACACCGCGCATAGCACCACCGGTGCA 3076
 1074 GlnThrCys---GlyIleAsnCys 1080
 3077 CAGAGCTGCATTTCCATTGATTGT 3100

RESULT 7

US-09-838-586-1
 ; Sequence 1, Application US/09838586
 ; Patent No. 6576818
 ; GENERAL INFORMATION:
 ; APPLICANT: Straker, D. et al.
 ; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
 ; TITLE OF INVENTION: Sequences
 ; FILE REFERENCE: 15621/02/US
 ; CURRENT APPLICATION NUMBER: US/09/838,586
 ; CURRENT FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/029,987
 ; PRIOR FILING DATE: 1996-10-29
 ; PRIOR APPLICATION NUMBER: 08/960,048

; PRIOR FILING DATE: 1997-10-29
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3328
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic oligonucleotide
 US-09-838-586-1

Alignment Scores:

Pred. No.: 0 Length: 3328
 Score: 3477.00 Matches: 657
 Percent Similarity: 73.60% Conservative: 129
 Best Local Similarity: 61.52% Mismatches: 176
 Query Match: 60.18% Indels: 106
 DB: 4 Gaps: 15

US-09-900-237a-30 (1-1080) x US-09-838-586-1 (1-3328)

Qy 18 ValCysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAla 37
 Db 200 GTTTGCCACACTTGTGGTGAACATGTTGGGTGAATGTTAATGTTGAACCTTTTGTGGCT 259
 Qy 38 CysAspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGly 57
 Db 260 TGGCATGAATGTAATTTCCCTATTGTAAGAGTTGTTTGAAGATGATCTTAAGGAAGGA 319
 Qy 58 ThrGlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIle 77
 Db 320 CGAAAGCTTGTTCCTTGT-----GATGATCCATAT--- 352
 Qy 78 ArgGlyGluGluGlyAspThrAspAlaAspGlySerAspPheAsnTyrProAla 97
 Db 353 -----GATGAAACCTGTTGGACGAT-----GTCGAG 379
 Qy 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsn 117
 Db 380 AAGGCCACCGCGCATCAATCGACNAATGGCTGCA-----CAT 415
 Qy 118 ThrGlyGlySerGlyAsnValGly---HisProLysTyrAspSerGlyGluIleGlyLeu 136
 Db 416 TTGACACAGTCTCAGGATGTTGGAAATTCATGCAGACATATCAGC-----AGTGTG 466
 Qy 137 SerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln 156
 Db 467 TCTACATTGGATAGT----- 481
 Qy 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
 Db 481 ----- 481
 Qy 177 IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe 196
 Db 482 -----GAAATG 487
 Qy 197 SerGlySerIleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLysGln 216
 Db 488 GCTGAACACCAATGGGAATTCGATTTGGAAGAACACAGGTTGGAAGTTCGAAGAAAGAAAG 547
 Qy 217 AspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAla 236
 Db 548 AACAAAGAAAGAGCGCTGCACAACT-----ARGGTT 580
 Qy 237 AlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGlu 256
 Db 581 GAAGAGAGGCTGAATCCACCTGAGCAACAATGGAAGATAAACCGGCACCGGATGCT 640
 Qy 257 ThrArgGlnProLysSerArgLysValProIleAlaSerSerIleAsnProTyrArg 276
 Db 641 TCC---CAGGCCCTCTCGATATATTCCATCCCGAAAGACAGACTTGCACCATACCGA 697

Qy 277 MetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsn 296
 Db 698 ACCGTGATCAATATGCGAATTGATCTTTTGGTCTTTTCTTCCATTATCGAGTAACAAAC 757
 Qy 297 ProValArgAsnAlaTyrProLeuTyrLeuSerValIleCysGluIleTyrPheAla 316
 Db 758 CCCGTTGCACAGTGTCTTTGGACTGTGGCTCAGTCATATGTGAATCTGGTTTGC 817
 Qy 317 LeuSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeu 336
 Db 818 TTTTCTCGGTGTGGATCAGTTCCCTAAGTGTATCTCTGTTAACAGGAAACATACAT 877
 Qy 337 AspArgLeuAlaLeuArgTyrAspArgGluGluProSerGlnLeuAlaValAsp 356
 Db 878 GACAGACTATCTGCAGATATGAAGAAGAGGTGAACCTGATGAACCTGTCGAGTTGAC 937
 Qy 357 IlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrVal 376
 Db 938 TTCTTCGTGAGTACAGTGGATCCATTGAAAGAGCTCCATTGATTACTGCCAATACTGTG 997
 Qy 377 LeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 396
 Db 998 CTTTCCATCTCTGCTTGGACTACCCGCTGGATAAGGTCTCTTCTTTATATATCTGATGAT 1057
 Qy 397 GlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaGluLysTrp 416
 Db 1058 GGTGCGCCCATGCTGACATTTGAATCTCTAGTAGAAACACGCGCACTTTGCAAGAAAGTGG 1117
 Qy 417 ValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGln 436
 Db 1118 GTTCCATCTCTCAAAAATTTTCCATTGAACCCCGGCGCACCTGAGTTTACTTCTCAG 1177
 Qy 437 LysIleAspTyrIleLysAspLysValGlnProSerPheValLysAspArgAlaMet 456
 Db 1178 AAGATTGATTACTTGAAGATAAAGTCAGCCCTCTTTTGTAAAGAACCTGAGACTATG 1237
 Qy 457 LysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLys 476
 Db 1238 AAAAGAGATTATGAAGAGATCAAAATTCGAATCAATCTTTAGTTGCAAGGCTCAGAAA 1297
 Qy 477 ValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArg 496
 Db 1298 ACACCTGATGAAGATGACATGCAAGATGGAATCTTTGGCCAGGAATAACCCCGCT 1357
 Qy 497 AspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly 516
 Db 1358 GATCACCTGGCATGATTGAGTTTCTCTGGATATAGTGTGCTCGTGACATCGAAGGA 1417
 Qy 517 AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 536
 Db 1418 AATGAATCTTCTCGACTGGTTTACGCTCTAGAGAGAAAGACCTGGCTACCAACACCA 1477
 Qy 537 LysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGln 556
 Db 1478 AAAAAGCTCGTGTGTAATAATGCTTTTGGTGTGCTGCTCAGTCTTCAAAAGCTCCC 1537
 Qy 557 TyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAla 576
 Db 1538 TTTCACTCTCAATCTTGATTGTGACCATATGTTAAACAATAGCAAGGAGTTAGGAGGCA 1597
 Qy 577 MetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGln 596
 Db 1598 ATGTGCTTCTTGATGAGCAACCAAGTTGGTCGAGATGTATGCTATGTCAGTTTCTCTCAA 1657
 Qy 597 ArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp 616
 Db 1658 AGATTGATGGCATAGTAGAGTATGATATGCCAATAGGAACACAGTTTCTTTGAT 1717
 Qy 617 IleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVal 636
 Db 1718 GTTAAACATGAAGGTCTTGATGGAATCCAAAGGCCAGTTTATGTGGAAACAGGTGTGT 1777
 Qy 637 PheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPhe 656

Db 1778 TTCAATAGCAGACACTTTATGGCTATGGTCCACCTTCAATGCCAAGTTTCCCAAGTCA 1837
Qy 657 LeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerAsp 676
Db 1838 TCCTCTCATCTTCGTGCTGCTGCTGCGCGGCAAGAACCTTAAAGATCCATCAGAG 1897
Qy 677 LysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIle 696
Db 1898 CTTTATAGGATGCAAAACCGGAAGAACTTGATGCTGCCATCTTTAACTTAGGAAAT 1957
Qy 697 GluGluGlyValGluGlyAlaGlyPheAspAsp---GluLysSerValLeuMetSerGln 715
Db 1958 GAC-----AATTATGATGATGATGAAGATCAATGTTGATCTCTCAA 1999
Qy 716 MetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGlu 735
Db 2000 ACAAGCTTTGAGAAACTTTTGGCTATCTTCAGCTCTTCANTGAATCTACACATAATGGAG 2059
Qy 736 TyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisVal 755
Db 2060 AATGAGAGTGGCTGAATCTGCCAACCTTCCACACTATCAAGAAAGCAATTCATGTC 2119
Qy 756 IleSerCysGlyTyrCysAlaPheLysSerGluTyrPglyThrGluIleGlyTrpIleTyrGly 775
Db 2120 ATCAGCTGTGCTATGAAGAGAACTGCTATGGGGAAGAGATGGATGGATATATGCT 2179
Qy 776 SerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerVal 795
Db 2180 TCAGTCACTGAGGATATCTTAACCGGCTTCAAAATGCACTGCCGAGGATGGAGATCGAT 2239
Qy 796 TyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 815
Db 2240 TACTGCATGCCCTTAAGCCACATTCAAAGATCTGCACCCATCAATCTGTCATCGG 2299
Qy 816 LeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCys 835
Db 2300 TTGCACCAAGTCTTCGATGGGCTCTTGGATCTGTGAAATTTCTAAGCAGGCAATGTC 2359
Qy 836 ProLeuTyrTyrGlyTyr---GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
Db 2360 CCTCATGATGATGGCTTGGAGGTGGTCTGCTTAAATGGCTTCAAGAGCTAGCATATATA 2419
Qy 855 AsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
Db 2420 AACACCATGCTATCTTTCATCTCCATCTTCACCTATGCTATGTTTCATACCAGCA 2479
Qy 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrp 894
Db 2480 ATCTGCTTCTCACAGGAAATTTATCATACCAACGCTCTCAAACTGGCAAGTGTCTTC 2539
Qy 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 914
Db 2540 TTTCTGGCCCTTTCTTCCATATATCGTACTGTGCTGCTTCTCGAGCTCCGATGGAGTGT 2599
Qy 915 ValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAla 934
Db 2600 GTACGATTTAGGACTTATGGCGTAACGAGAGTTTTGGTTCATCGGTGGCGCTTCAGCC 2659
Qy 935 HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
Db 2660 CATCTCTTGGCGCTTCCCAAGGTTTCTTAAGATGCTTGGCGGCAATTCACACCACTTT 2719
Qy 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
Db 2720 ACTGTCACGCAAAAGCAGCTGAT---GATGAGATTTTGGGAGCTCTACATTTGGAAA 2776
Qy 975 TrpThrThrLeuLeuIleProThrThrIleLeuIleIleAsnMetValGlyValVal 994
Db 2777 TGGACTACACTTCTAATCCCTCCAAACACACTCTCATCTGTCACATGTTGGTGTGCT 2836
Qy 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLys 1014

Db 2837 GCCGATTTCTCCGATCCCTCAACAAAGGTACGAAGCTTGGGACCACCTCTTTGGCAAA 2896
Qy 1015 LeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
Db 2897 GGTGTTCTTTCTCTCTGGGTCACTCCATCTTTATCCATCTCTCAAGGCTTTATGGGA 2956
Qy 1035 ArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePhe 1054
Db 2957 CGCCAAACAGACACCAACCATTTGTTGCTTGGTCAGTGTGTTGGCTTCTGTCTTC 3016
Qy 1055 SerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle--- 1073
Db 3017 TCTCTGTTTGGGTTCCGATCAACCCGTTTGTGACACCGCGCATAGCACCCGCTGTCA 3076
Qy 1074 GlnThrCys---GlyIleAsnCys 1080
Db 3077 CAGAGCTGCATTTCCATTGATGT 3100
RESULT 8
US-09-221-013A-1
; Sequence 1, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1, 4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2248
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1887)
US-09-221-013A-1
Alignment Scores:
Pred. No.: 2,36e-273 Length: 2248
Score: 2629.50 Matches: 468
Percent Similarity: 87.46% Conservative: 76
Best Local Similarity: 75.24% Mismatches: 71
Query Match: 45.51% Indels: 7
DB: 4 Gaps: 4
US-09-900-237A-30 (1-1080) x US-09-221-013A-1 (1-2248)
Qy 454 ArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLys 473
Db 1 CGAGCTATGAGAGAGAGATGATGAAGAGTTTAAAGTGAGATAAATGCTCTTGTGGCAAA 60
Qy 474 AlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsn 493
Db 61 GCACAGAAAATCCCTGGAGNAGGCTGGACAATGAGGATGGTACTCCCTCGGCTCGTAAC 120
Qy 494 AsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyClyLeuAsp 513
Db 121 AACACTAGAGATCATCTCGGAATGATACAGGTGTTCTTAGGCCATAGTGGGGCTCTGGAT 180
Qy 514 ThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPhe 533
Db 181 ACCGATGGAATGAGCTGCCTAGACTCATCTATGTTCTCGTGAAGAGCGGCTGGATTT 240
Qy 534 GlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThr 553

Db 241 CAACACACAAAAGGCTGGAGCTATGAATGCATCGATCCGTGTATCTGCTGTTCTTACC 300
Qy 554 AsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaVal 573
Db 301 AATGGAGCATATCTTTTGAACGTGGATGTGATCATTACTTAAATACAGTAAGGCTATT 360
Qy 574 ArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGln 593
Db 361 AAAGAAGCTATGTGTTTCATGATGGACCCCGCTATTGGAAGAAGTGTGCTATGTCCAG 420
Qy 594 PheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrVal 613
Db 421 TTCCTCAGCGTTTGACGTTATGATTTGCACGATCGATATGCCACAGAAATATAGTC 480
Qy 614 PhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThr 633
Db 481 TTTTTCGATATTAACATGAAGGGGTGGATGGTATCCACGGTCAGTATATGTGGGTACT 540
Qy 634 GlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys--- 652
Db 541 GGTGTGTTGTTTAAATAGCGAGCTCTATATGGGTATGATCCCTGTTTTCGCGAAGAAGAT 600
Qy 653 ---LysProGlyPheLeuAla---SerLeuCysGlyGlyLysLysLysAlaSerLysSer 670
Db 601 TTAGAACCAATATATTGTCAAGAGCTGTTCGGGTCAAGGAAGAAGGTAAAGTAGC 660
Qy 671 LysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSerValProVal 690
Db 661 AAGAAGTATATACTACGAAAAGAGAGAGGATCAACACAGAGTCACTCCAAATGCTCCACTT 720
Qy 691 PheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGlyLysSer 710
Db 721 TTCAATATGAGGACATCGATGAGGGTTTGA-----GGTTATGATGATGAGAGGTCT 774
Qy 711 ValLeuMetSerGlnMetSerLeuGlnLysArgPheGlyGlnSerAlaAlaPheValAla 730
Db 775 ATTCATATGCCAGAGAGGTAGAGAGGTTTGGTCAGTCGCCGTATTATTATGGC 834
Qy 731 SerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLys 750
Db 835 GCACCTTCATGGAACAAAGGGCGCATTCACCAACCAACCAATCCGCTACTCTTCTGAAG 894
Qy 751 GluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIle 770
Db 895 GAGGCTATTATGTTATAGCTGTGGTACAGACAAAGACTGAATGGGGCAAAAGAGATT 954
Qy 771 GlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArg 790
Db 955 GGTGGATCTATGTTCCGTGACGAAGATATTCTTACTGGTTCAAGATCATGCCCGG 1014
Qy 791 GlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIle 810
Db 1015 GGTGGATATCGATCTACTGCAATCCTCCACGCCCTGGTTCAAGGGATCTGCCCAATC 1074
Qy 811 AsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleu 830
Db 1075 AATCTTCTGATCGTTGAAACCAAGTTCTTCGATGGGCTTTGGGATCTATCGAGATTCTT 1134
Qy 831 PheSerArgHisCysProLeuTyrPheGlyTyrGlyArgLeuLysPheLeuGluArg 850
Db 1135 CTTAGCAGACATTGCTCTATCTGGTATGTTACCATGGAAGTTGAGACTTTTGAGAGG 1194
Qy 851 PheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCys 870
Db 1195 ATCGCTTATATCAACACCATCGTCTATCTCTATCATCCATCCCTCTATTATGGGTATTGT 1254
Qy 871 IleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeu 890
Db 1255 ATTCCTCCCGTGTGTTGCTCATCCACGACAGATTTCATCATCCGAGATGAAGCACTAC 1314
Qy 891 AlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMet 910

Db 1315 GCGAGTATTGGTTTCATTCTACTCTTCATCTCAATTGCTGTGACTGGAATCCTGAAACTG 1374
Qy 911 ArgTyrSerGlyValGlyIleAspGluTyrTyrPheArgAsnGluGlnPheTyrValIleGly 930
Db 1375 AAATGGAAACGTTGTGAGCATTTGAGATTGGTGGAGGAACAACCAAGTTCTGGGTCAATTGGT 1434
Qy 931 GlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIle 950
Db 1435 GGCATATCCACCAATCTTTTGTGCTCTTCAAGGTCTACTTAAAGGTCTTGTCTGGTATC 1494
Qy 951 AspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeu 970
Db 1495 AACACCAACTTCACCGTTTACATCTTAAAGCCACAAACAAATGGGATTTTGCAAACTC 1554
Qy 971 TyrMetPheLysTyrThrThrLeuLeuLeuProProThrThrIleLeuIleAsnMet 990
Db 1555 TACATCTTCAATGGACAGCTCTTCTCATTCACCAACCAACCGCTCTACTTGTGAACCTC 1614
Qy 991 ValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyPro 1010
Db 1615 ATAGGCAATTTGGCTGTGTCTCTTATGCTGTAAACAGTGGCTACCACTGCTGGGTCCG 1674
Qy 1011 LeuPheGlyLysLeuPheAlaPheTyrValIleValHisLeuTyrProPheLeuLys 1030
Db 1675 CTTTTCGGGAAGCTCTTCTTCGCTTATGGGTTATTCGCCATCTCTACCCCTTCTTGA 1734
Qy 1031 GlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeu 1050
Db 1735 GGTCTGTGGTGGAGACAAACCGCAACCAACCAATCGTCATGTCTGTCTTCTTCTC 1794
Qy 1051 AlaSerIlePheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGly 1070
Db 1795 GCCTCCATCTTCGTGCTTTGGTGGTCAAGATCAATCCCTTGTGGAC-----GCCAAT 1848
Qy 1071 ProAsn 1072
Db 1849 CCAAT 1854

RESULT 9
US-09-221-013A-4
; Sequence 4, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(5009)
; OTHER INFORMATION: N is A, T, G or C.
US-09-221-013A-4

Alignment Scores:
Pred. No.: 5,25e-263 Length: 5009
Score: 2539.50 Matches: 559
Percent Similarity: 56.71% Conservative: 130
Best Local Similarity: 46.01% Mismatches: 166
Query Match: 43.95% Indels: 362


```
Db 3491 ATTTTGTAAATGACAGGTGTTCTTTGGGTCATAGTGGAGTTTCGTGATACGGATGCTAATGA 3550
Qy 518 uLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLy 538
Db 3551 GTTACCACGCTCTAGTGTATGTTCTCGTGAGAGCGGCTGGATTGATCACCAAGAA 3610
Qy 538 sAlaGlyAlaMetAsnAlaLeuVal----- 546
Db 3611 AGCTGGAGCTATGAATTCCTTGTTGTAAGTATTAATGTGTTCTTTTATTAATGATCTCTCT 3670
Qy 547 -----ArgVa 548
Db 3671 TTCGGAGCCCTGACTTCTCATATAAACTAAACTCATCTTACTTCTTCTTGAAGATCCGAGT 3730
Qy 548 lSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAs 568
Db 3731 CTCTGCTGTCTTATCAAAACGCTCTTACCTTCTTAATGTGCAATGTGATCACTACATCAA 3790
Qy 568 nAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGl 588
Db 3791 CAACAGCAACCAATTAGAGNACTATGTGTTCATGATGACCCGCAATCCGGAAAGAA 3850
Qy 588 nValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAl 608
Db 3851 AGTTTGTATTGCTTCAGTTTCCGCAGAGATTGATGGGATTGATAGACATGATAGATACTC 3910
Qy 608 aAsnArgAsnThrValPhePheAsp----- 616
Db 3911 AAACCGTAACGTTGTGTTCTTTGATGATGTGCTCTTATCTCTTTTGTGTTTCTGTT 3970
Qy 617 -----IleA 618
Db 3971 TATGTTTTAGTCTTTTCCCTCTTTCTCAITTTGATATGTTTTCGTGTGAAACAGATTA 4030
Qy 618 snLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheA 638
Db 4031 ACATGAAGGTCTTGATGGGATACAGAGCCGATATATGTCGGACAGGTGTGTGTGTTA 4090
Qy 638 snArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuA 658
Db 4091 GAAACAGGCTCTTTATGGTTTGTATGCACCAAGAAAGAAACCAACCCAGGCAAAACCT 4150
Qy 658 laSer-----LeuCysGlyGlyLysLysLysAlaSerLysSerL 671
Db 4151 GTAACCTTGGCCTAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4204
Qy 671 yLysArgSerSerAspLysLysSerAsnLysHisValAspSerSerValProValP 691
Db 4205 AAACGAAAGCCAAAGATAAGAAACTAACACTAAA-----GAGACTTCAAAGCAGATTC 4258
Qy 691 heAsnLeuGluAspIleGluGluGlyVal-----GluGlyAla--- 703
Db 4259 ATGCGCTAGAGAAATGTCGAGGAAGTGTATTCGTCGCCAGGTAAATAAAGAAAGAAAAA 4318
Qy 704 -----GlyPhe----- 705
Db 4319 AAACATTTCTATTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4378
Qy 706 -----AspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGly 723
Db 4379 GTGTCAAAATGTTGAGAAGATCTGAAGCAACACAAATTTGAAATTTGGAGAAAGTTTGG 4438
Qy 724 GlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSer 743
Db 4439 CAATCTCCGGTTTTCGTTGCTCTGCTGTTCTACAGAACCGTGGAGTTCCCGCTTAACGCA 4498
Qy 744 ThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 763
Db 4499 AGCCCCGATGTTTGTATAGAGAGCCATTCAGTTTATAGTCGCGGTACGAGATATAA 4558
Qy 764 SerGluTrpGly----- 767
Db 4559 ACCGAATGGGAAAGAGGTAGAAAAACATTACAAAGTTTTTCAACTTCTGAAAACTAGAA 4618
```

```
Qy 768 -----ThrGluIleGlyTrpIleTyrGlySe 776
Db 4619 AAGTTCTTGTGATCTCAATTCCTGCTGATAATCACACGACGATCGGTGGATTTATGGATC 4678
Qy 776 rValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTy 796
Db 4679 GGTGACTGAAGATATCTGACGGGTTTCAAGATGCAATGCCATGATGGAGATCTGTGTA 4738
Qy 796 rCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLe 816
Db 4739 CTGTATGCTTAAGCGTTCAGCTTTTAAAGGATCTGCTCTTAACTATTGTCAAGATCGTCT 4798
Qy 816 uAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysPr 836
Db 4799 TCATCAAGTTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTCCTTGAGCAGACATTGTCC 4858
Qy 836 oLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnTh 856
Db 4859 GATATGATGTTATGTTGGTGTGTTTAAATGGTTGGAGAGATCTCTTACATCAATC 4918
Qy 856 rThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCy 876
Db 4919 TGTGCTATCTCTGGACTTCACCTTCATTCATGATGCTCTATTGTTCTCTCCCGGGTTTG 4978
Qy 876 sLeuLeuThrGlyLysPheIleMetProGlu 886
Db 4979 TTTACTCACAGSAAATTCATCGTCCCTGAG 5009
RESULT 10
US-08-960-048-2
; Sequence 2, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4612
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-2
Alignment Scores:
Pred. No.: 9,71e-187 Length: 4612
Score: 1833.00 Matches: 510
Percent Similarity: 42.72% Conservative: 80
Best Local Similarity: 36.93% Mismatches: 78
Query Match: 31.72% Indels: 714
DB: 3 Gaps: 29
US-09-900-237A-30 (1-1080) x US-08-960-048-2 (1-4612)
```

```
Qy 413 AlaArgLysTrpValPropheValLysTyr-Asp----- 424
Db 39 GCGAGGAGATGGGTTCCGTTTGTAAAGACATTGATCACCTAGGGGGCCGACGTCTCT 98
Qy 425 -----IleGluProArgAl 429
Db 99 AAGCGGTGCTGCTCTCTTATCCCAAGCAAAACATTCTTCTGTTAATGTTGAGCCCAAGGC 158
Qy 429 aProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGln----- 446
Db 159 GCCCGAGTTTATTTCAATGAGAGAAATTGATTATTTGAAGGACCAAGGTCCATTATTAAC 218
```

446 ----- 446
Db TCGGGTCCGGCGCTCAAAATAAAGTTACTCTCTTAATAATAAACTTCCTGTCACAG 278
447 --ProSerPheValIysAspArgAlaMetLysArgGluTyrGluGluPheLysIleA 466
Db TACCTAGCTTTGTTAAAGAACCGAGAGCCATGAAAGGGAATATGAAGAAATTTAAAGTAA 338
446 rGileAsnAla----- 469
Db GGATCAATGCATCGATCGAAACAATTTCTTGCTCTCGGTACTTTTCCCTTATACTTCTT 398
470 -----LeuValSerLysAlaLeuLysValProGluGluGlyTyr 482
Db AAAATTCATTCCTAGTTAGTATAGTACAAAGCTCAGAGAAACCAAGAGAGATGG 458
483 IleMetGlnAspGlyThrProTyrProGly-Asn----- 493
Db GTGATGCAAGATGGCACCCCATGGCCCGGAAATCATCGTTTTCGAGTCTCTTTGGTCT 518
494 -----AsnThrArgAspHisPr 499
Db TCTTCTACCCACTAGTTCTACCGTGGGGTACCGGGCTTTTAAACACTCGTATCATCC 578
499 ocGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly- 516
Db TGGATGATTAGGTCTATCTAGGAAGTGCCTGGTGCATCGATGTGGATGGCAATGTGAG 638
516 ----- 516
639 CACTAGTAGGACCTTACTAAGTCCAGATAGATCTTACCGGCCACGTGAGCTACACCTAC 698
517 --AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisH 536
Db CGAAAGAGCTGCTCGACTGTGTCTATGTTCTCGTGAGAAACGACCTGGTTATCAGCACC 758
536 isLysLysAla----- 539
Db ATAGAAAGCCGTTTCTCGACGGAGCTGNAACAGATACAAAGAGCACTCTTTGCTGGACCA 818
540 -----Gly-AlaMetAsnAlaLeuValArgValSerAlaValLeu 552
Db ATAGTCGTGGTATCTTTCGGCGTGTGAGAAATGCTCTGTTTCGAGTTTCTCGAGTGTCT 878
553 ThrAsnGlyGlnTyrMetLeuAsnLeuAspCys----- 563
Db ACTAATGCACCTTCATATGAATCTGGATTGCAAGCACTCTTACGAGACCAAGCTCAAAG 938
564 -----AspHisTyrIleAsnAs 569
Db ACGTCACGNATGATACGTGGGAAGTATAACTTAGACCTAATCTGATCATATACACAA 998
569 nSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyPro----- 587
Db TAGCAAGGCCATGAGGAGAGCGATGCTTTTAAATGGATCCTCAGTTTGGAACTAGTAA 1058
587 ----- 587
1059 TGATGTTGTTATCGTTCCGGTACTCCCTTCGTGTACACGAAAAATTTACCTAGGAGTCAAAAC 1118
588 -----Gln-ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspA 606
Db CTAAGAGCTTTGTTATGTTCAATTTCCACAGAGATTTGATGGTATGATCGTCATGATC 1178
606 rGlyTyrAlaAsn----- 609
Db GATATGCTAATCTTCTTCCAAACAATAACAAGTTAAAGGTGTCTCTAAACTACCATAACTA 1238
610 -----Arg-AsnThrValPhePheAspIleAsnLeuArgGlyLeu 622
Db GCAGTACTAGCTATACGATTAGGAATGTTGCTTCTTTGATATCAACATGTTGGGATTA 1298

623 AspGlyIleGlnGlyProValTyrValGlyThr----- 633
Db GATGGACTTCAAGGCCCTGTATATGTAGGCACCTTTACAACAGAAAAAATACTAGTTGTA 1358
634 -----GlyCysValPheAsnAr 639
1359 CAACCTAATCTACCTGAAGTTCGGGACATATACATCCGTGAGGCTGTGTTTCAACAG 1418
639 gThrAlaIleTyrGlyTyrGluProIleLysAlaLysLysPro----- 654
Db CGAGCATTTGATGGCTACGATCCACCACTCTCTGAGAAACGACCAAGATGTCCCAAC 1478
654 ----- 654
1479 AAAAGTTGTCGCTCCGTAAACATACCGATGTAGTGGTTCAGAGACTCTTTGCTGCTTCT 1538
655 -----GlyPheLeuAla---SerLeu-CysGlyGlyLysLysLysAlaAs 668
Db ACACATGTGATGTGCGGCTTCTTGGTGTCTGTTGTGCGAGGTTCTAGGAAGAAAT 1598
668 erLysSerLysLysArgSerAspLysLysLysSerAsnLysHisVal----- 684
Db CAAGAGAAAGTGTACACTAACGACCGAGAACCAACACACACAA-CGCTTCCAAAG 1657
684 ----- 684
1658 ATCCCTTCTTCTTCTTCTTTCGTGAAAAAGAGGCTTACTCGGAGGCTTTTATACGG 1717
684 ----- 684
1718 AAAAAAGAGAAGATGATGGGCAAAACATATGTCATCTTTTCCGGAATGAGCCTCCAG 1777
685 -----AspSerSerValP 689
1778 AAAATATGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1837
689 roValPheAsnLeuGluAspIleGluGluGlyValGluGly----- 702
Db CAGTCTTTGATCTCGAAGAAATCGAAGAGGCTTGAAGGATACGAAAGAAATGCTTTT 1897
703 -----AlaGlyPheAsp- 706
1898 CCCAGACGTGGTFCAGAACTAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1957
707 AspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAla 726
1958 AACGAGAAATCCACATTAATGTCGCAAGAAATTCGAGAAACGATTCGACAAATCACCG 2017
727 AlapheValAlaSer----- 731
2018 GTTTTCAVTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2077
732 -----ThrLeuMetGluTyrGlyGlyValProGlnSerSe 743
2078 TCTTAGTGCCCAAAAGTAAACGACCACTTTGATGAAATGTTGGCTTCTTCTTCTTCTTCT 2137
743 rThrProGluSerLeuLeuLysGluAlaIleHis----- 754
2138 TAATTCACATCCTGATTAAGAGGCCATTCAGTTGAACTACCTTTTACCACCGGAAG 2197
755 -----ValIleSerCysGlyT 760
2198 GACTTCTTGTAAAGTGTAGTACTAATTTCTCCGTAAGTTCGTAATTAGCTGTGTT 2257
760 rGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerVal----- 777
2258 ATGAGAAAAAATCAGTGGGGCAAGAGATCGATGGAATTTATGGTCTGGTGGCAATTAA 2317
777 ----- 777
2318 TCGACACCAATCTTCTTTTGTGACTCACCCGCTTCTCTAGCTTACCTAAATACCCAGC 2377
778 ---ThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrTrpArgSerValTyr 796

```
Db 2378 CACACGGAAGATATATTAAACAGGTTTCAAGATGCAATTGTAGAGGGTGGAAATCGTTTAT 2437
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 797 CysMetProLysArg----- 801
Db 2438 TGTGTACCGAAATGCTCTATATAATTGTCAAAGTTCTACGTAACATCTCCACACTT 2497
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 802 -----ProAlaPheLysGlySerAlaProIleAsnLeuSe 813
Db 2498 TAGCCAAATTAACACATGCTTTTACCGGCATTCAAAGGGTCCGCTCCAATCAATCTCTC 2557
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 813 rAspArgLeuAsnGlnValLeuArgTTPAlaLeu----- 824
Db 2558 GGATCGGTTCACCAAGTTTGTAGATGGCACTCTGGCCGCTAAAGTTTCCAGCGGAGGTT 2617
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 825 -----GlySerValGluLeuL 830
Db 2618 AGTTAGAGAGCTAGCAACAGTGGTTCAAAACTCTACCCGCTGATGGTTCTGTAGAAATTT 2677
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 830 euPheSerArgHisCysProLeuTrpTyrglyTyrglyGlyArgLeuLys----- 846
Db 2678 TCCTTAGTCGCTACTGTCCTTGTATGTTGTAAGTGGAAACTGAAATGGACCAAGA 2737
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 847 -----P 847
Db 2738 CATCTTTAAAGGAATCAGCAGTGACAGGTGAACACCATACCAATPACCACCTTTTGACTTT 2797
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 847 he-LeuGluArgPheAlaTyrlleAsnThrTrpIleTyrrProLeuThrSerLeuProLeu 866
Db 2798 ACCCTCAGAGGCTTGTATTATCAACACCATTTGTTACCCCTTTTCACTCGATCCCTTTA 2857
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 867 LeuValTyrcys----- 870
Db 2858 CTCGCTATTGTAGAGCTCTCCGAACGAATAGTTGTGTAAAGTGGAAATGGAAAGTGGAG 2917
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 -----Ile-LeuProAlaIleCysLeuLeuThrGlyLysPheIl 883
Db 2918 CTAGGGAATGAGCGGATAACATCTATTCCAGCTGTTTGCTTCTCACCAGCAATTCAT 2977
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 883 eMetProGluIleSerAsnLeuAlaSerIleTrp----- 894
Db 2978 CATTCCAACTCTAAGCAACCTTACAAGTGTGTGATAGGTGCGACAAACAGAAAGTGGC 3037
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 895 -----PheIleAlaLeuPheL 900
Db 3038 CGTTTAAGTAGTAGGTTGAGATTCTGTTGGAATGTTCAACACAGTCTTGACACTTTTCC 3097
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 900 euSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu- 919
Db 3098 TCTCCATCATTTGCAACTGGAGTGCTTGAACCTCGATGGAGCGGGTTAGCATCCAAGAAC 3157
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 919 ----- 919
Db 3158 CGTGAAGAGGAGGTAGTAACGTTGACCTCAGCAACTTGAAGTACTCGCCCAATCG 3217
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 920 -----TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeu 936
Db 3218 TAGCAAGACTGGTGGCCCAATGAACAATCTGGGTGATCGAGGTGTCGCGCCACTTT 3277
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 937 PheAlaValPhe----- 940
Db 3278 TTTGCTGCTCCGTTCTGACCACCGGCTTACTTGTGTAAGACCACCTAGCCTCCACAGAG 3337
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 941 -----GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAs 953
Db 3338 GCGGGTAGAAAAACACAGAGAAGGCGCTCTCTCAAGTCTTAGCTGGAGTAGACACCAA 3397
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 953 nPheThrValThrSerLysAlaAsnAspGluGlu-Gly----- 965
Db 3398 CTTACCGGTAAACAGAAAGCAGCAGATACTCTCCGGAGGAGTTTCAAGATCGACCTC 3457
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 966 -----AspPheAlaGluLeuT 971
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 3458 ATCTGTGGTTGAAGTGGCACTGTGTTTTTCGTGCTGTGCTATGAGAATTCGGTGAACCTTT 3517
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 971 yrMetPheLysTrpThrThrLeuLeuLeuProProThrThrThrLeuLeuLeu----- 988
Db 3518 ATCTCTCAAAATGACAACTCTCTTATATCCCTCCCACTCTGATAATACTGTCTTTAAG 3577
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 988 ----- 988
Db 3578 CCACCTGAATAGAGAAGTTTACCTGTTGAGAGAAATTAGGAGGGTTCGAGACTATTAT 3637
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 989 ----AsnMetValGlyValValAlaGlyThrSerTyrrAlaIleAsnSerGlyTyrrGlnSer 1007
Db 3638 GACAAACATGTCGGAGTCGTGGCGGAGTTTCAGACGCAATCAACAACGCGCTATGTTCA 3697
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1008 TrpGlyProLeuPhe----- 1012
Db 3698 TGGGGTCCATTGTTTGTACCAGGCTCAGACCGGCTCAAAGTCTGCGTTAGTTGTTGCC 3757
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1013 -----GlyLysLeuPhePheAlaPheTrpValIleValHi 1024
Db 3758 GATACCAAGTACCCAGGTAAACATCGGCAAACTGTCTTCGCATTCTGGGTCACTCTTCA 3817
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1024 sLeuTyrrProPheLeuLysGlyLeuMetGlyArg----- 1035
Db 3818 TCCTTACCCATCTCTCAAAGGTTTGATGGGAGAGCGGTTTGACAAGAAGCGTAAGACCC 3877
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1036 -----GlnAsnArgThrProT 1041
Db 3878 AGTAAGAAGTAGAAAATGGGTAAAGAGTTTCCAAACTACCCCTCACAACACAGGACGCCA 3937
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1041 hrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTrp----- 1058
Db 3938 CCATTGTTGCTTGTGTCTACTCTTGGCATCGAATTTCTCACTGTTGGTGGTTTG 3997
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1058 ----- 1058
Db 3998 TCCTCGGGTGTAAACAACAGAACCCAGGTATGAAAACCGTAGTAAAAGAGTGACCAA 4057
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1059 ----ValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGly 1077
Db 4058 ACGTACGGATGATCCCTTCTTGCCCAAAACAAACAGGTCCAGTCTTAAACATATGTGC 4117
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1078 IleAsnCys 1080
Db 4118 GTGAGTGC 4126
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 11
US-09-838-586-2
; Sequence 2, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4612
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-838-586-2
Alignment Scores: 9.71e-187 Length: 4612
Pred. No.: 9.71e-187 Length: 4612
```

Score: 1833.00 Matches: 510
Percent Similarity: 42.72% Conservative: 80
Best Local Similarity: 36.93% Mismatches: 78
Query Match: 31.72% Indels: 714
DB: 4 Gaps: 29

US-09-900-237A-30 (1-1080) x US-09-838-586-2 (1-4612)

QY 413 AlaArgLysTyrValProPheValLysLysTyr-Asp----- 424
DB 39 GCGAGGAGATGGGTTCCGTTTTTGTAAAGACATTCATCACCTAGGGGGCCGACGCTCCTT 98
QY 425 -----IleGluProArgAl 429
DB 99 AAGCGTGTCTCCTCTACCAAGGCAAAACATTTCTTCGTTAATGTAGGCCCGAGGCG 158
QY 429 aProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGln----- 446
DB 159 GCCCGAGTTTATTTCAATGAGAAGATTGATTATTTGAAGGCAAGGTCATATTACAAC 218
QY 446 ----- 446
DB 219 TCGGTCGCGGGCCCTCAAAATAAAGTTACTCTTCTAACTAATAAACTTCCTGTTCCAGG 278
QY 447 --ProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLysIleA 466
DB 279 TACCTAGCTTTGTTAAAGAACGAGAGCCCATGAAAGGAAATATGAAGAAATTTAAAGTAA 338
QY 466 rGileAsnAla----- 469
DB 339 GGATCAATGCATGGATCGAAACAATTTCTTGCTCTCGGTACTTTTCCCTTATACTTCTT 398
QY 470 -----LeuValSerLysAlaLeuLysValProGluGluGlyTyr 482
DB 399 AAATTTCAATCTAGTTAGTATAGTAAAGTCTCAGAAACACAGAAAGATGG 458
QY 483 IleMetGlnAspGlyThrProTyrProGly-Asn----- 493
DB 459 GTGATGCAAGATGGCACCCCATGGCCCGGAAATCATCGTTTCGAGTCTCTCTTGGTCT 518
QY 494 -----AsnThrArgAspHisPr 499
DB 519 TCTTCTACCCACTAGTTCTACCGTGGGTACCGGGCCTTTTAACTACCTCGTATCATCC 578
QY 499 oGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly----- 516
DB 579 TGGAAATGATTCAAGTCTATCAGGAAGTCCCGGTGCATCGATGCGGATGGCAATGTGAG 638
QY 516 ----- 516
DB 639 CACTAGTAGACCTTACTAAGTCCAGATAGATCTTCCAGGCCACGCTGAGCTACACCTAC 698
QY 517 --AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisH 536
DB 699 CGAAAGAGCTGCTCGACTGTCTATGTCTTCTCGTGAAGAACGACCTGGTTATCAGCAC 758
QY 536 IsLysLysAla----- 539
DB 759 ATAAGAAAGCCGTTTCTCGACGGAGCTGAACAGATACAAGAGCACTCTTTGCTGGACCA 818
QY 540 -----Gly-AlaMetAsnAlaLeuValArgValSerAlaValLeu 552
DB 819 ATAGTCGTGGTATCTTTTCGGCGTGTGAGAAATGCTCTGGTTCGAGTTTCTCGAGTGT 878
QY 553 ThrAsnGlyGlnTyrMetLeuAsnLeuAspCys----- 563
DB 879 ACTAATGCACCCCTTCATATTGAATCTGGATTGCAGCTCTTTACGAGCACAAGCTCAAG 938
QY 564 -----AspHisTyrIleAsnAs 569
DB 939 ACGTCACGAATGATTACGTGGGAAGTATACTTAGACCTAACTGATCATTTACATCAACAA 998
QY 569 nSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyPro----- 587

DB 999 TAGCAAGCCCATGAGGAAGCATGTGCTTTTAAATGGATCCTCAGTTTGGAACTAGTAA 1058
QY 587 ----- 587
DB 1059 TGTAGTTATTATCGTTCGCTACTCCCTTCGTACACGAAATAATACCTAGGAGTCAAC 1118
QY 588 ----Gln-ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspA 606
DB 1119 CTAAGAAGCTTTGTTATGTTCATTTCCACAGAGATTTGATGGTATTGATCGTCATGATC 1178
QY 606 rGlyrAlaAsn----- 609
DB 1179 GATAGCTTAATCTTCTCGAAACAATACAAGTTAAAGGTGTCTCTAAACTACCATAACTA 1238
QY 610 -----Arg-AsnThrValPhePheAspIleAsnLeuArgGlyLeu 622
DB 1239 CGAGTACTAGCTATACGATTAGGAAATGTTGCTCTTTTGATATCAACATGTTGGATTA 1298
QY 623 AspGlyIleGlnGlyProValTyrValGlyThr----- 633
DB 1299 GATGGACTTCAAGGCCCTGTATATGTAGGCACCTTTACAACAGAGAATACTAGTTGTA 1358
QY 634 -----GlyCysValPheAsnAr 639
DB 1359 CAACCCTAATCTACCTGAAGTTCCGGGACATATACATCCGTGAGGCTGTGTTTCAACAG 1418
QY 639 gThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysPro----- 654
DB 1419 CGAGCATTTGATGGCTAGCTACGATCCACAGTCTCTCAGAAACGACCAGAGATGCCACAC 1478
QY 654 ----- 654
DB 1479 AAAAGTTGTCGTCGTAACATACCGATGTAGGTGGTCAGAGACTTTTGTGCTTTCT 1538
QY 655 -----GlyPheLeuAla---SerLeu-CysGlyGlyLysLysLysAlaAs 668
DB 1539 ACACATGTGATTGCTGGCCTTCTTGTGTTGCTGTGTTGTCGAGGCTCTTAGAAGAAAT 1598
QY 668 erLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisVal----- 684
DB 1599 CAAGAAGAAAGTGTACACTACGACCGGAGAACCAACACGACACAA-CCCTCCCAAG 1657
QY 684 ----- 684
DB 1658 ATCCCTTCTTAGTTTCTTCTTTCGTAAGAAAGGCTTACTCGGAGGCTTTTATACGG 1717
QY 684 ----- 684
DB 1718 AAAAAGAAGAAGATGATGGGCAAAACTATGTCTTCTTCCGAATGAGCCTCCAG 1777
QY 685 -----AspSerSerValP 689
DB 1778 AAATATAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1837
QY 689 roValPheAsnLeuGluAspIleGluGluGlyValGluGly----- 702
DB 1838 CAGTCTTTGATCTCGAAGAAATCGAAGAGGCTTGAAGGATACGAAGAATTTGCTTTT 1897
QY 703 -----AlaGlyPheAsp- 706
DB 1898 CCCAGACGTGCTCAGAAACTAGAGCTTCTTTAGCTTCTTCCGAACTTCTTATGTTCTT 1957
QY 707 AspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAla 726
DB 1958 AACGAGAATCGACATTATGTCGAGAGAAATTTTCGAGAACGATTCGAGAACATCACCG 2017
QY 727 AlaPheValAlaSer----- 731
DB 2018 GTTTTCAITGCTCTCTTTTAGCTGTAATTACAGCGCTCTTCTTAAAGCTCTTTGCTAAGCC 2077
QY 732 -----ThrLeuMetGluTyrGlyGlyValProGlnSerSe 743


```

; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1741
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1741)
US-09-221-013A-13

Alignment Scores:
Pred. No.: 3,11e-179 Length: 1741
Score: 1756.50 Matches: 346
Percent Similarity: 73.16% Conservative: 66
Best Local Similarity: 61.46% Mismatches: 100
Query Match: 30.40% Indels: 51
DB: 4 Gaps: 13

US-09-900-237A-30 (1-1080) x US-09-221-013A-13 (1-1741)
QY 2 AspGlyAspAla---AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB 167 GACGGCGACCGCCCGCGCTAAGCAGGAGAGTGAATGCTCAGTCTGCCAG 226
QY 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB 227 ATTGTGGCGACACTGTTGGCGTCTCGCCACCGCGACGCTCTTGTTCCTGCAATGAG 286
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
DB 287 TCGCCCTTCCTCGGCTCGCCGCTTCTACGATACGAAACGAAAGGAGGAAACCGATGC 346
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB 347 TGCCCCCATGCAAGACTAGATACAAAGCAGCAAAAGGTGCGCTAGAGTTCAGGCGAT 406
QY 81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrPro----- 96
DB 407 GAG---GAAGAAGAGATGTTGATGACCTGGACATGAATTCATTATAAGCATGGCAAT 463
QY 97 AlaSerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMet 116
DB 464 GGCNAAGGTCAGAG-----TGCCAGATA 487
QY 117 AsnThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeu 136
DB 488 CAGACACAGGGGAA-----GATGTTACCTGCTTCATCT 523
QY 137 SerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln 156
DB 524 TCTCGCCACCAACACATCGGATTCCTCCGT-----CTGCAAGTGGGCAACAG 571
QY 157 MetSerGlyGluIleProGlyAlaSerProCysPheHisHisMetMetSerProThrGlyAsn 176
DB 572 ATCTCAGGAGATCCCTGATGCTCCCGCGATTCGCCATTCCTATC----- 616
QY 177 IleSerArgArgAlaProPheProTyrValAsnHisSer----- 189
DB 617 -----CGCAGCGGAACATCAAGCTATGTTGATCCCAAGTCTCCAGTCTCTGAGGAT 670
QY 190 ProAsnProSerArgGluPhe-----SerGlySerIleGlyAsnValAlaTrpLysGluArg 208
DB 671 GTGACCCCTCCAAAGGACTTGAATTCCTATGGGATTAACAGAGTTTAAAGTGTGACTGGCAAGAAAGA 730

```

```

QY 209 ValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSer 228
DB 731 GTTGCCAGCTGGAGGAACAAGCAGGACAAATAATATGTCAGGTAGCTAAT----- 781
QY 229 IleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMet 248
DB 782 ---AAATATCCAGAGGCAAGGGGGA---GACATGGAA---GGGACTGGTTCAAAATGCT 832
QY 249 GluAspAlaLeuLeuAsnAspGluThrArgGlnProLysSerArgLysValProIleAla 268
DB 833 GAAGATATCCAAATGGTTCATGATCAGCTCTACCTCTGAGCGCATAGTGCCTATCCCT 892
QY 269 SerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIle 288
DB 893 TCAAAACCACTCAACCTTTACCGAATGTTATCATCTCCGCTCTATCATCTCATGTTTC 952
QY 289 PheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSer 308
DB 953 TTCTTCCAATATCGTGTCACTCATCCAGTGGGATGCTTATGGATTGTGGCTAGTATCT 1012
QY 309 ValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPhe 328
DB 1013 GTTATCTGTGAAATTTGGTTGCTTATCTCGCTCTAGATCAATTCCTCAAGTGGTAC 1072
QY 329 ProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGlu 348
DB 1073 CCGATAAACCGTGAACATACCTTGACAGGCTTGATTTGAGATATGATAGGAGGGAGAG 1132
QY 349 ProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluPro 368
DB 1133 CCATCACAGCTTGCTCCCATTCATGCTTTGTGTCAGTACGTGGATCCACTAAAGGAACCT 1192
QY 369 ProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys 388
DB 1193 CCTGTGATCACAGCAAACTGTTTGTTCATCTCTGGCTGTGGATTACCTCTGTGACAAA 1252
QY 389 ValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGlu 408
DB 1253 GGTGATGCTATGTTCTGACCATGGTTCAGCTATGTTAACTTTTGAGGCTCTGTGAGAA 1312
QY 409 ThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArg 428
DB 1313 ACTGCAGAAATTTGCTAGGAAGTGGGTTCCGTTTTCAGAAAGCACAATATTGAACCCGA 1372
QY 429 AlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSer 448
DB 1373 GCTCCAGAGTTTACTTTGCTCAAAAATAGATTACCTGAAGAGGCTGGACCATGCTGTC 1432
QY 449 PheValLysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsn 468
DB 1433 TTTGTTAAAGAAAGCGGCAATGAGAGAGATGAGAGATTAAGAAATTCAGGTACGGATCAAT 1492
QY 469 AlaLeuValSerLysAlaLeuLysValProGluGluGlyTyrIleMetGlnAspGlyThr 488
DB 1493 GCTCTTTGTCGAAGGCACAAAAGTACCTGAAGAGGCTGGACCATGCTGTCGTCACCT 1552
QY 489 ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 508
DB 1553 GGTGTGCTTGGGAATAACCAAGGGATACCTCGGCATGATTACAGGTGTTCTTGGGGCAC 1612
QY 509 SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 528
DB 1613 AGTGGTGGGCTTGACACTGATGGTAACAGAGTTGCCACGGCTGTCTACGCTCTCGTGAA 1672
QY 529 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuValArgVal 548
DB 1673 AAGAGCCAGGATTCAGCATCACAAAGAGGCTGTGCAATGAATGAATGATTGCTGTGTA 1732
QY 549 SerAlaVal 551
DB 1733 TCTGCTGTG 1741

```



```

RESULT 13
US-09-313-294A-6485
; Sequence 6485, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6485
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351882H1
US-09-313-294A-6485
Alignment Scores:
Pred. No.: 1.6e-42 Length: 281
Score: 478.00 Matches: 88
Percent Similarity: 98.92% Conservativeness: 4
Best Local Similarity: 94.62% Mismatches: 1
Query Match: 8.27% Indels: 0
DB: Gaps: 0

US-09-900-237A-30 (1-1080) x US-09-313-294A-6485 (1-281)

QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro 873
Db 2 ATCAACACCAACATCTACCGCGTCACGTCCTCCGCTCCTCATTTACTGTATCTGCCT 61

QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
Db 62 GCACATCTCTGTCACGGGAGTTTATATCCAGAGATCAACAACTTCGCTAGTATC 121

QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913
Db 122 TGGTTCATCTCTCTTCATCTCGATCTTCGCCAGGGGTATCTCGAGATGAGTGGAGC 181

QY 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTyrValIleGlyIleSer 933
Db 182 GCGCTGGGCATCGACGAGTGTGGAGGAACGAGCAGTCTCTGGTTCATCGGAGGCATCTCC 241

QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysVal 946
Db 242 GCCACCTCTTCGCCGCTTCACGGGCTCTCTCAAGGTG 280

RESULT 14
US-09-313-294A-4753
; Sequence 4753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4753
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350285H1
US-09-313-294A-4753
Alignment Scores:
Pred. No.: 1.32e-38 Length: 291
Score: 442.00 Matches: 81
Percent Similarity: 89.69% Conservativeness: 6
Best Local Similarity: 83.51% Mismatches: 9
Query Match: 7.65% Indels: 1
DB: Gaps: 0

US-09-900-237A-30 (1-1080) x US-09-313-294A-4753 (1-291)

QY 790 ArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaPro 809

```

```

; OTHER INFORMATION: Incyte ID No. 6476212 700348982H1
; NAME/KEY: unsure
; LOCATION: 64, 189
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4753
Alignment Scores:
Pred. No.: 2.3e-42 Length: 300
Score: 477.00 Matches: 92
Percent Similarity: 97.00% Conservativeness: 5
Best Local Similarity: 92.00% Mismatches: 3
Query Match: 8.26% Indels: 1
DB: Gaps: 0

US-09-900-237A-30 (1-1080) x US-09-313-294A-4753 (1-300)

QY 886 GluIleSerAsnLeuAlaLeuSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThr 905
Db 2 CAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTTCATCTCGATCTTCGCCACG 61

QY 906 GlyIleLeuGluMetArgTyrTrpSerGlyValGlyIleAspGluTyrTrpArgAsnGluGln 925
Db 62 GGNATCTCGAGATGAGTGTGGAGCGGGTGGGCATCGACGAGTGTGGAGGAGACGAGCAG 121

QY 926 PheTyrValIleGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLys 945
Db 122 TTCGGGTGATCGGGGCATCTCCGCGCATCTTCGCCGCTTCACGGGCTCTCTCAAG 181

QY 946 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGly 965
Db 182 GTGCTGGCGCATCGACACCACTTCACCGTCACCTCCAGGCTCGGACGAGGAGCGGC 241

QY 966 AspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuLeuProThrThrIle 985
Db 242 GACTTCGCGGA-CTGTACATGTTCAAGTGGACGACGCTCTTATCCGCCGCCACCAACCATC 300

RESULT 15
US-09-313-294A-5519
; Sequence 5519, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5519
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350285H1
US-09-313-294A-5519
Alignment Scores:
Pred. No.: 1.32e-38 Length: 291
Score: 442.00 Matches: 81
Percent Similarity: 89.69% Conservativeness: 6
Best Local Similarity: 83.51% Mismatches: 9
Query Match: 7.65% Indels: 1
DB: Gaps: 0

US-09-900-237A-30 (1-1080) x US-09-313-294A-5519 (1-291)

QY 790 ArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaPro 809

```

Db	3	CGANGCTGGCGATCAATCTACTGCATGCCTAAGCGACNAGCTTTCAAGGGANCTGCTCCT	62
Qy	810	- IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIle	829
Db	63	ATNCNACCTTTCCGATCGTTTGAATCAAGTCTTCGCTGGGCTCTTGGTTCATATGSAAT	122
Qy	829	eLeuPheSerArgHisCysProLeuTrpTyvGlyTyrcrGlyArgLeuLysPheLeuG1	849
Db	123	TCITTTTCAAGCAGGCANTTCTCCCATATGGTATGGCTATGGAGTTCGGCTAAATTCCTGGA	182
Qy	849	uArgPheAlaTyrlleAsnThrThrlleTyvProLeuThrSerLeuProLeuLeuValTy	869
Db	183	GAGATTTGCTTATATCANCACAAACAATTTATCCACTCANATCAACTCNGCTCCTCNTGTGA	242
Qy	869	rCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetPro	885
Db	243	CTGCATATGGCAGCAGATNTGCTTCTCACTGGGAAGTTTCATCATCCCA	291

Search completed: August 22, 2004, 22:25:26
Job time : 300 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 20:43:24 ; Search time 1097 Seconds
(without alignments)
4836.011 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MGDADALSKGRGAGDVQ.....VDFPTTLAGPNITQGINC 1080

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delxext 7.0

Searched: 3228839 seqs, 245606551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0900237/runat_18082004_081516_14735/app_query.fasta_1.1223
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US0900237@cgn2_1_1_750@runat_18082004_081516_14735
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=130 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
------------	-------------	--------	-------	-------------

ALIGNMENTS

RESULT 1

US-09-900-237-29
; Sequence 29, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 3626
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-900-237-29

Alignment Scores:

```

Pred. No.:      0      Length:      3626
Score:      5778.00      Matches:      1080
Percent Similarity:      100.00%      Conservative:      0
Best local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      9      Gaps:      0

US-09-900-237A-30 (1-1080) x US-09-900-237-29 (1-3626)

QY      1      MetAspGlyAspAlaAspAlaLeuIysSerGlyArgHisGlyAlaGlyAspValCysGln      20
DB      60      ATGACGGCGACGCGACGCCCTGAAGTCCGGAGGACACGGGGCGGGAGCGTGTGCCAG      119

QY      21      IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal      40
DB      120      ATCTCGCGCGACGGGCTGGGACCCACGTTGGACGGCGAGCTCTTACCGCCCTGCCACGTC      179

QY      41      CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla      60
DB      180      TGCCGCTTCCCGTCTGCCGCCCTCTGTACGAGCAGCAGCGCAAGGGGCGACCCAGGCC      239

QY      61      CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu      80
DB      240      TGCCCTCCAGTGCAGACCAAGTACAAGCGCCACAGAGGAGGCCAGCGATCCGCGGGAG      299

QY      81      GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr      100
DB      300      GAAGGCGACGACACTGATGCCGATGATGTAGTACTTCAACTACCCCTGCATCTGGCACT      359

QY      101      GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly      120
DB      360      GAGGACCAAGACAGCAAGATTGCTGACAGGATGCCAGCTGGCGCATGACACACCGGGGC      419

QY      121      SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp      140
DB      420      AGTGGCAATTGTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCTCTCCAAGTATGAC      479

QY      141      SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu      160
DB      480      AGTGAGAGATCCCTTAGGGGATACCTCCCTTCAGTCACCAACAGCACCATGTCAGAGAA      539

QY      161      IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg      180
DB      540      ATCCCTGGAGCTTCGCTGATCATCATGATGATGTCCTCCCTACGGGAAACATCAGCAGAGC      599

QY      181      AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle      200
DB      600      GCTCCGTTTCCCTATGTGAATCATTCACCAATCCGTCGAAGGAGTTCCTCCGGCAGTATT      659

QY      201      GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla      220
DB      660      GGGAAATGTGCTGGAAAGAGAGAGATTGATGGCTGGAAATGAACGACGACAAAGGGTGG      719

QY      221      IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle      240
DB      720      ATTCCTCATGACTAATGGACAAAGCATTCCTCCCTCTGAAGGTCCGGGCGAGCTACTGACATC      779

QY      241      AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro      260
DB      780      GATGCATCTACTGAAATACAACATGGAAGACGCTTTTACTGAATGATGAATCGCCAGCCT      839

QY      261      LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal      280
DB      840      CTATCTAGAAAGTCCCATTTGCTTCCCAAAATPAATTCCTCAGAAATGTCTAATGTT      899

QY      281      LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn      300
DB      900      CTGCGTTTGGTTGTTCTTAAGCATCTTCCTGCACTACCGTCTCACAAATCCTGTGCGTAAT      959

QY      301      AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle      320
DB      1000      TGTGGGGGCAAGAAAGGCAAGCAAGTCAAGAAAGAGAGCTCAGATAAGAAAAAGTCG      2099

```

```

DB      960      GCATACCCACTGTGGCTTTTATCTGTTATATGTGAGATTGGTTGCTTTATCTCGGATA      1019

QY      321      LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrIleuAspArgLeuAla      340
DB      1020      CTGGATCAGTTTCCGAAGTGGTTTCCAATCAACCGGAGACCTACTCTTGATAGACTGGCT      1079

QY      341      LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer      360
DB      1080      TTAAGGTATGACCGAGAGGTGAACGCTCTCAGTTGGCTGCTGTGACATATTTGTCAGT      1139

QY      361      ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu      380
DB      1140      ACAGTCGACCCCTTGAAGGAGCCACTATCGTCACTGCGCAACACTGTGTCATCCATCTT      1199

QY      381      AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyIleSerMet      400
DB      1200      GCTGTGTATTATCCCGTGACAAAGGTCTCTGCTATGTATCTGTATGACGAGGACTTCAATG      1259

QY      401      LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaAlaArgLysTrpValProPheVal      420
DB      1260      CTGACTTTTGAGGCATTGGCTGAGACTTCAGAGTTTGTAGGAAATGGGTACCATTTGTG      1319

QY      421      LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr      440
DB      1320      AAGAAGTATGACATTTGAACCCAGAGCTCCGAGTTTATCTTTTGGCAGAAAAATTGATTAC      1379

QY      441      LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr      460
DB      1380      CTGAAGACAAAGTCCAGCCTTATTGTTAAAGACCCCGGGCCCATGAAGAGAGAAATAT      1439

QY      461      GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu      480
DB      1440      GAAGAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAGTCCCCGAGGAA      1499

QY      481      GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly      500
DB      1500      GGATGGATCATCAAGATGGCACACCATCGCCGAGGAAACAATACACAGGATCATCTGGA      1559

QY      501      MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro      520
DB      1560      ATGATTCAGGTTTCCCTTGGTCACAGTGGTGGCTGTGATCTAGGGTAATAGGTCTCCC      1619

QY      521      ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly      540
DB      1620      CGTTTATTTATGTCTCGTGAAGAGCTCTCGGTTCCAGCACCAAGAGGCTGTG      1679

QY      541      AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn      560
DB      1680      GCCATGAATGCCCTTGTCTCGTCTCTCAGCTGTCTTACTTAATGGACAATACATGTTGAAT      1739

QY      561      LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu      580
DB      1740      CTTGATTGTGATCACTACATCAACACAGAGGCTGTCCGAGAGAGCTATGTGCTTCCTA      1799

QY      581      MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly      600
DB      1800      ATGGATCAACACCTTAGGTCGCAAGTCTGTATGTGCAAGTCCCAACAAAGGTTTGTATGG      1859

QY      601      IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg      620
DB      1860      ATTGATGAATATGATCGATATGCAACAGGAAACACTGTCTTTTGTATTAATTAACCTTAGG      1919

QY      621      GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr      640
DB      1920      GGCCTTGACGCATTCAGGACCGATTATGTGGGAACCTGGTTGTGTTTTTCAACAGAAACA      1979

QY      641      AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu      660
DB      1980      GCTATCTATGTTATGAGCCCCCAATTAAGCGGAAGAACCCAGGTTTCTTGGCATCACTA      2039

QY      661      CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer      680
DB      2040      TGTGGGGGCAAGAAAGGCAAGCAAGTCAAGAAAGAGAGCTCAGATAAGAAAAAGTCG      2099

```

```
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2100 AACAGCATGTGGACAGTCTGTGTCCAGTATTCAATCTCGAAGACATAGAGGAGGGTGT 2159
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2160 GAAGTGCTGGGTTTGATGATGAGAAATCAGTCTCATGTCTCAATGAGCTTAGAGAAG 2219
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2220 AGATTTGGCCAGTCAGCAGCATTTGTTCCTCCACTCTGATGGAATATGGTGTTCCT 2279
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2280 CAGTCTCCACTCCAGAAATCTCTTTGAAGAAGCTATCCATGTCATGATCTGTGGGTAT 2339
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db 2340 GAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAAAT 2399
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrPheArgSerValTyrCysMetProLys 800
Db 2400 ATTCTTACTCGATTCAAGATGCGACGCAAGAGGCTGGCGTTTCAGTCTATTGCATGCCCAAG 2459
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2460 CGCCCACTTTCAGGGATCTGCCCATCAATCTTTCAGATCTGTGAACCAAGTCTG 2519
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2520 CGGTGGGCTCTCGGTCTCTGTGAAATCTTTTCAGCCGGCATTCGCCCTTATGGTATGGC 2579
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2580 TACGGAGGGCGCCTCAAGTTCTCTGGAGAGATTGCTTTACATCAACACCACCATTTTACCCA 2639
QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2640 CTAACCTCTCTCCCGCTTCTAGTCTATGTATATTGCTGTCTGTCTGTCTCAGTCTGCA 2699
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2700 AAGTTCAATCATGCCAGCATTAGCAACTTGGCCAGTATCTGTTTCATTGGCTTCTTCTT 2759
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 2760 TCAATTTTCGCCACTCGGTATCTCTAGATGAGGTGAGTGTGTGGCAATTCACGAGTGG 2819
QY 921 TyrArgAsnGluGlnPheThrValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2820 TGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACATCTGTGTGGCGTCTTT 2879
QY 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 2880 CAGGTCTTCTGAAGTGCTTGGCGGTATCGACACCAACTTCACCTGTCACTCAAGGCT 2939
QY 961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle 980
Db 2940 AATGACGAAGAGGCGACTTTGTCTGAGCTCTACATGTTCAGTGGACGAGCTTCTCATC 2999
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValAlaGlyThrSerTyrAla 1000
Db 3000 CCTCCGACGACCAATTTTGATCAATTAACATGGTTGGTGTGGTGTGGCAGCTCTCAGCC 3059
QY 1001 IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyr 1020
Db 3060 ATCAACAGTGGTACCAATCATGGGGCGGCTCTTTGGGAGCTCTTCTTTGGCTTCTGG 3119
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3120 GTGATTGTTCACTTATACCATTCCTCAAGGGTCTTATGGGCGAGGCAAAACCGCACACCG 3179
```

```
QY 1041 ThrIleValIleValTyrPheAlaValLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Db 3180 ACATTTGTCTATCTCTTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3239
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3240 GTTGATCCATTCACTACCGGTCTCGGTGGCCCAATATCCAAACCTGTGGCATCAACATCG 3299

RESULT 2
US-10-437-963-72402
; Sequence 72402, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72402
; LENGTH: 3962
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72784C.1
US-10-437-963-72402

Alignment Scores:
Pred. No.: 0 Length: 3962
Score: 5452.50 Matches: 999
Percent Similarity: 97.22% Conservative: 52
Best Local Similarity: 92.41% Mismatches: 29
Query Match: 94.37% Indels: 1
DB: 17 Gaps: 1

US-09-900-237A-30 (1-1080) x US-10-437-963-72402 (1-3962)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 373 ATGACCGGCACCGGATGCGTGAAGTCGGGAGGACCGGAGCGCGGCGAGCGTGCAG 432
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 433 ATCTGGCGGCGCGGCGTGGGCGACGACGCGGAGGCGACGTGTTCGCCGCTCGCAGCTC 492
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 493 TGGCGCTTCCGCTGTGCGCCCTGCTACGAGTACGAGCGCAAGGATGCGACCCAGGCT 552
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 553 TGCCCCCAGTGCAGAACCAAGTACAAAGCGCCAAAGGGGAGCCCGCGATCCCGTGGGGAG 612
QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 613 GAAGGGCGAGATACTGATGCTGATGATGTGTCAGTACATACCACTACCTGATCTGGCAGT 672
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db 673 GCGGACCAAGAGCAGAGATGCTGATAGGATGGCAGTTGGCGCATGATGCTGGGGGT 732
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 733 GGTGGAGACGTGGCGCGCTCCCAAGTAGTAGTACAGTGGCGAGATCGGGCTCACCAAGATGAC 792
```



```
Db 2953 CCATTGACATCGATCCCGCTTCTCATATACTGTGTGTTTGGCTGTCTATCTGTGTTGCTCACT 3012
Qy 880 GtlylvsPheIleMetProGluIleSerAsnLeuAlaSerIleTrrPheIleAlaLeuPhe 899
Db 3013 GGGAGATTTCATCCCGAGATTAGCAACTTTGCTAGTATTGGTTTCATCTCTCTCTTC 3072
Qy 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrrPserGlyValGlyIleAspGlu 919
Db 3073 ATTTCATTTTGGCCATGGTATCTTGAGATGAGTGGATGGTGGTGGATCGATGAG 3132
Qy 920 TrrPrrArgAsnGluInPheTrrPvalIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 3133 TGGTGGAGGAATGAACACTTCTGGGTATTGAGGATATCTGGCATCTTTTGGCCGTC 3192
Qy 940 PheGlnGlyLeuLeuLysValIleAlaGlyIleAspThrAsnPheThrValThrSerIys 959
Db 3193 TTCAGGGTCTCCCTCAAGGTGCTTGCTGGTATCGACCAATTTCACTGTCACTCAAG 3252
Qy 960 AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrrPheThrLeuLeu 979
Db 3253 GCTTCTGATGAAGTGGGACTTTGCTGAGCTCTACATGTTCAAGTGGACACACGCTTCTC 3312
Qy 980 IleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3313 ATCCACCGGACGACCATCTTGATCATTAACCTGGTTCGGTGTGTTGCTGTGATCTCATAC 3372
Qy 1000 AlaIleAsnSerGlyTrrGlnSerTrrGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3373 GCGATCAACAGCGGCTACCATCAATGAGGACCGCTCTTTGGCAAGCTCTCTTTGGCCPTC 3432
Qy 1020 TrrValIleValHisLeuTyrPrrPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3433 TGGGTGATTGTCACCTTGATCCCTTCTCAAGGTCTTATGGGTTCGCAAAACCGCACT 3492
Qy 1040 ProThrIleValIleValTrrPAlaValLeuLeuAlaSerIlePheSerLeuLeuTrrPVal 1059
Db 3493 CCGACCATCGTGTGTTGTTGGCAATCTTCTGGCTTCGATCTCTCATTCGCTGGGTT 3552
Qy 1060 ArgValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3553 CGCATCATCTCAATCACACCGGTGTACCGGCCGAGATACCCAAACATGTGGCATCAAC 3612
Qy 1080 Cys 1080
Db 3613 TGC 3615

RESULT 3
US-10-627-132-21
; Sequence 21, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; PRIOR APPLICATION NUMBER: 2003-07-25
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 3799
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: 3757-3775, 3777, 3782
; OTHER INFORMATION: n = A,T,C or G
US-10-627-132-21

Alignment Scores:
Pred. No.: 0 Length: 3799
Score: 5423.50 Matches: 996
Percent Similarity: 96.48% Conservative: 46
Best Local Similarity: 92.22% Mismatches: 37
Query Match: 93.86% Indels: 1
DB: 13 Gaps: 1

US-09-900-237A-30 (1-1080) x US-10-627-132-21 (1-3799)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 238 ATGGAGGGCGACGCGGACCGCGTGAAGTCCGGGAGGCGCGGGGAGGCGAGGTGTGCCAG 297
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 298 ATCTGCGGCGATGGCGTGGGCACTACGCGGAGGAGGAGACGTCTTCAACGCTTGGACGTC 357
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlyThrGlnAla 60
Db 358 TGGCGGTTCCGGTGTGCGGCCCTGCTACGAGTACGAGCGCAAGGACGGCACCAAGCG 417
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 418 TGCCCCCATGTCAAAAACAAAGTCAAGCGCCACAAAGGGAGTCCAGCGATCCGAGGGAG 477
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerClyThr 100
Db 478 GAAGGAGACGATACTGATGCGGATGCTAGCGACTTCAACTACCTGCACTCGGCAAT 537
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrrArgMetAsnThrGlyGly 120
Db 538 GACGACCAAGACGAGAGATTTGTCAGAGATGCGAGCTGGCGCATGAATGCTGGGGGC 597
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 598 AGCGGGGATGTTGGCCGCCGCCCAAGTATGACAGTGGTGAGATCGGGCTTACCAAGTACGAC 657
Qy 141 SerGlyGluIleProArgGlyTrrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 658 AGTGGTGAGATCCCTCGGGGATACATCCCGTCACTCAACACGCGAGATTTTCGGAGAA 717
Qy 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 718 ATCCCTGGTCTCCCTCGACCATCATATGATGTCTCTACTGGGAAACATTTGGCAGGCGC 777
Qy 181 AlaProPheProTrrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 778 GCCCATTTCCCTATATGATCATTCATCAATCCGTCAAGGAAATTTCTCTGGTAGCGTT 837
Qy 201 GlyAsnValAlaTrrLysGluArgValAspGlyTrrLysMetLysGlnAspLysGlyAla 220
Db 838 GGGAAATGTTGCTCGAAAGAGAGGGTTGATGGCTTGGAAAATGCAAGCAGGACCAAGGAA 897
Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 898 ATTCCTCATGAGATGGCACCAAGCATTTGCTCCCTCTGAGGGCCGGGGTGTGGTATATT 957
Qy 241 AspAlaSerThrGluTrrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 958 GATGCATCAACTGATTAACAATGGAAGATGCTCTATTAAACGATGAACTCGCCAGCCT 1017
Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTrrArgMetValIleVal 280
Db 1018 CTATCTAGGAAAGTTCCACTTCTCCAGGATAAATCATACAGGATGGTCAATTGTG 1077
Qy 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
```


Db	1078	CTACGATTGATTGTTCTTAAGCATCTTCTTGCACTACCGGATCACAAATCCTGTGCGTAAT	1137
QY	301	AlaTyrProLeuTyrLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle	320
Db	1138	GCATACCCACGTGTGCTTCTATCTGTATATGTAGATCTGTTGCTCTTTCTCGGATA	1197
QY	321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
Db	1198	TTGGATCAGTTTCCAAAGTGGTTTCCATCAACCCGGAGACTTACTTGTATAGATCGCA	1257
QY	341	LeuArgTyrAspArgGluGluProSerGlnLeuAlaValAspIlePheValSer	360
Db	1258	TTAAGGTATGACCGGAAAGTGAGCATCTCAGTTGGCTGCTGTGACATTTTGTCAGT	1317
QY	361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380
Db	1318	ACTGTGCAACCAATGAAGAGGCTCTCTTGTTCATCTGCAATACCGTGTATCCATTCNC	1377
QY	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1378	GCTGTGGACTATCCTGTGGATAAGGTCCTTGTCTATGTATCTGTATGATGAGCTGCTATG	1437
QY	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
Db	1438	CTCACATTTGATGCACTAGCTAGACTTCAGAGTTTGTCTAGAAAATGGGTGCCATTTGTT	1497
QY	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1498	AAGAAGTACCAATTAAGGCTCTAGAGCTCTGAATGGTACTTCTCCAGAAAATGATTAC	1557
QY	441	LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr	460
Db	1558	TTGAAGGACAAAGTGACACCTTCATTTGTTAAAGACCGCGGCGCATGAAGAGAAATAT	1617
QY	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1618	GAAGAATTCAAAATTAAGGCTTGTGTTAAGGCACAAAAAGTCCTCGAGGAA	1677
QY	481	GlyTyrIleMetGlnAspGlyThrProTyrProGlyValAsnThrArgAspHisProGly	500
Db	1678	GGATGGATCATGACAGATGGCACCATGGCCAGGAAACAATACCAGGGACCATCTCGGA	1737
QY	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1738	ATGATTCAGTTTTCCTTGGTCACAGTGGTGTCTTGATCTAGGGTAATGAGTACCC	1797
QY	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly	540
Db	1798	CGTTTGGTCTATGTTTCTCGTGAACAAACGTCCTGGATTCCAGCATCACAAAGAAAGCTG	1857
QY	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560
Db	1858	GCATGAATGCTCTTGTCCGCGTCTCAGCTGTGCTTACCATGGACATACATGTTGAAT	1917
QY	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580
Db	1918	CTTGATTGTGATCACTACATCAACAACAGTAAGGCTCTCAGGGAAGCTATGTTCCCT	1977
QY	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600
Db	1978	ATGGATTCCTAACTAGGAGGAGTGCTGCTATGTTTCACTGTTTCCCCAGAGGTTGCGAT	2037
QY	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620
Db	2038	ATTGATAGGAATGATCGATATGCCAACAGAAACACCGTGTGTTTTCGATATTAACTTGGA	2097
QY	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640
Db	2098	GGTCTTGATGGCATCCAAAGACAGTTTATGTGGGCACCTGGGTGTTTTCACAGAAACA	2157
QY	641	AlaIleTyrGlyTyrGluProProLysAlaLysLysProGlyPheLeuAlaSerLeu	660
Db	2158	GCTCTATATGGTTATGAGCCCCCAATTAAAGCAAAAGAGGGTGGTTTCTTGTCATCACTA	2217
QY	661	CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer	680
Db	2218	TGTGTGCGAGGAAGGAGCAATCAAGAAG--GGCTCAGACAAGAAAAGTCA	2274
QY	681	AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal	700
Db	2275	CAGAAGCATGTGACAGTTCTGTGCCAGTATTCAATCTTTGAAGATATAGAGGAGGAGTT	2334
QY	701	GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys	720
Db	2335	GAAGCGCTGGATTTGATGATGAGAAATCACTTCTATGTCTCAATGAGCTTGGAGAAG	2394
QY	721	ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro	740
Db	2395	AGATTTTGCCAAATCTGACAGCTTTGTTCGATCCACTCTGATGGAATATGGTGGTCTCT	2454
QY	741	GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr	760
Db	2455	CAGTCTGGGACTCCAGAATCTCTTCTGAAAGAAAGCTATCCATGTATCAAGTTGTGGCTAC	2514
QY	761	GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp	780
Db	2515	GAGGACAAAGATTGAATGGGAACCTGAGATTGGGTGGATCTATGGTCTCTGACGGAGAT	2574
QY	781	IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys	800
Db	2575	ATTCTCACTGGTTCAAGATGCACGACGAGGCTGGCGGTGCGATCTACTTGATGCTTAAG	2634
QY	801	ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu	820
Db	2635	CGCGCGGCTTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCCAGGTGCTC	2694
QY	821	ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly	840
Db	2695	CGGTGGGTCTCGGTTCAGTGAAATPCTTTTCAGCGGCAATGCCCTATGGTACGGG	2754
QY	841	TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	860
Db	2755	TACGAGGACCGCTGAGTTCTTGGAGAGATTCGCTACATCAACACCATCTACCCG	2814
QY	861	LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly	880
Db	2815	CTCAGTCCCTCCGCTCCTCACTTACTGTATCTGCTGCTGCCATCTGCCGTGTACAGGG	2874
QY	881	LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu	900
Db	2875	AAAGTTCAATCATCCAGAGATCAGCAACTTCGCTAGTATCTGGTTTCACTCTCTTTCATC	2934
QY	901	SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr	920
Db	2935	TCGATCTTCGCCACGGTATCTCGAGATGAGTGGAGCGCGCTGGGCATCGACGAGTGG	2994
QY	921	TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe	940
Db	2995	TGGAGGAACAGCAGTTCTGGGTTCATCGGAGCATCTCCGCCACCTCTTCGCGCTCTC	3054
QY	941	GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla	960
Db	3055	CAGGGCTCTCCTCAAGGTGTTGCCGATCGACACCACTTCCCGTCACTTCCAAAGGCC	3114
QY	961	AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle	980
Db	3115	TCGGATGAAGACGGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGACACTTCTGATC	3174
QY	981	ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaIleThrSerTyrAla	1000
Db	3175	CGCGCCACCATCTGATCATCAACCTTGGTGGCGTGTGTGGCGGCATCTCTTACGCC	3234
QY	1001	IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp	1020
Db	3235	ATCAACAGCGGGTACAGTCTGTTGGGTCTGCGCTCTTCGGCAAGCTCTTCTTCGCCCTCT	3294

Qy	461	GlulPheLysIleArgIleAenAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1618	GAAGAAATTCAAATTAGGGTAATAGGCTTGTTGCTAAGGCACAAAAAGTCCCTGAGGAA	1677
Qy	481	GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	500
Db	1678	GGATGGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCGAGGACCAATCTCTGGGA	1737
Qy	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1738	ATGATTCAGGTTTTCCTTGGTCACAGTGGTGGTCTTGATACTGAGGGTATGAGCTACCC	1797
Qy	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly	540
Db	1798	CGTTTGGTCTATGTTCTCGTGA AAAACGCTCTGGATTCAGCATCAACAAGAAAGCTGGT	1857
Qy	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560
Db	1858	GCCATGAATGCTCTGTGCCGGCTCTACGTGTGCTTACCAATGACAAATACATGTTGAAT	1917
Qy	561	LeuAspCysAspHisTyrIleAAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580
Db	1918	CTTGATTGTGATCACTACATCAACAACAGTAAAGCTCTCAGGGAAGCTATGTGCTTCCTT	1977
Qy	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600
Db	1978	ATGGATCCTTAACCTAGGAAGAGTGTCTGCTATGTTCAAGTTTCCCCAGAGGTTTCGATGGT	2037
Qy	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620
Db	2038	ATTGATAGGAATGATTCATATGCCACAGGAACACCGTGTGTTTTTCGATATTAACCTTGAGA	2097
Qy	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640
Db	2098	GGTCTTGATGSCATCCAAGGACCAGTTTATGTGGGCATCGCTGTGTTTTCAACAGAACCA	2157
Qy	641	AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu	660
Db	2158	GCTCTATATGGTTATGAGCCCCCAATTAAGCAAAAGAGGGTGGTGTTCCTGTGTCATCACTA	2217
Qy	661	CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer	680
Db	2218	TGTGGTGGCAGAAAGAGGAGCAATCAAGAG--GGCTCAGACACAGAAAAAGTCA	2274
Qy	681	AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal	700
Db	2275	CAGAAGCATGTGGACAGTTCTGTGCCAGTATTCATCTTGAAGATA TAGAGGAGGAGTT	2334
Qy	701	GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys	720
Db	2335	GAAGGCGCTGGATTGTATGATGAGAAATCACTCTTATGTCTCAATAGAGCTTGGAGAAG	2394
Qy	721	ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro	740
Db	2395	AGATTGGCCAACTCTGCAGCTTTTGTGCGTCCACTCTGATGGAAATATGTTGGTGTTCCT	2454
Qy	741	GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr	760
Db	2455	CAGTCTGCGACTCCAGAATCTCTTCTGAAAGAGCTATCCATGTCATAAGTTGTGGCTAC	2514
Qy	761	GluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAsp	780
Db	2515	GAGGACAAGATTGAATGGGAACTGAGATTGGGTGGATCTATGTTGTTCTGTGACGGAAGAT	2574
Qy	781	IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys	800
Db	2575	ATTCCTACTGGGTTCAAGATGCACGACGAGCGTGGCGTGCATCTACTGCATGCCCTAAG	2634
Qy	801	ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu	820
Db	2635	CGGCCGCGCTTCAAGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCGAGTGCCTC	2694

Qy	821	ArgTnpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuThrTyrGly	840
Db	2695	CGGTGGGCTCTCGTTTCAGTGGAATCCCTTTTCAGCCGGCATTCGCCCTATGGTACGGG	2754
Qy	841	TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	860
Db	2755	TACGGAGGACGCCTGAAGTTCCTTGGAGAGATTCCCTACATCAACACCACCATCTACCG	2814
Qy	861	LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly	880
Db	2815	CTCAGGTCCCTCCGCTCCCTCAATTACTGTATCTCTGCCTGCCTCTGCCTCACGGGG	2874
Qy	881	LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu	900
Db	2875	AAGTTCAATCATCCAGAGATCACAACTTCGCTAGTATCTGTGTTCATCTCTCTCTCATC	2934
Qy	901	SerIlePheAlaThrGlyIleLeuGluMetArgTnpSerGlyValGlyIleAspGluTrp	920
Db	2935	TCGATCTTCGCCACGGGTATCCTTGGAGATGAGGTGGAGCGCGTGGGCATCGACGAGTGG	2994
Qy	921	TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe	940
Db	2995	TGGAGNACAGACAGTTCGTGGGTCAATCGGAGGCATCTCCGCCCACTCTTCGCCGCTTC	3054
Qy	941	GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla	960
Db	3055	CAGGGCTCTCCAAGGTGCTTGC CGGCATCGACACCAACTTCACCGTCACCTCCAAAGCC	3114
Qy	961	AsnAspGluGlyAspPheAlaGluLeuTyrMetPheIlystrPhrThrLeuLeuIle	980
Db	3115	TCGGATGAACAGCGGCATCTCGCGAGCTGTACATGTTCAAGTGGACGACACTTCGTATC	3174
Qy	981	ProProThrThrIleLeuIleIleAsnMetValGlyValAlaGlyThrSertYrAla	1000
Db	3175	CCGCCACACCATCTCGATCATCACTGCTCGCGGTGTTCGCCGCATCTCTACGCC	3234
Qy	1001	IleAsnSerGlyTyrGlnSerTnpGlyProLeuPheGlyLysLeuPheAlaPheTrp	1020
Db	3235	ATCAACAGCGGTACAGTCTGGGGTCCGCTCTTCGGCAAGCTCTTCTTCGCTTCTTG	3294
Qy	1021	ValIleValHisIeuTyrProPheLeuLysGlyLeuMetClyArgGlnAsnArgThrPro	1040
Db	3295	GTGATCGTTCACCTGTACCGCTTCCTCAAGGGGTCTCATGGGTGGGAGAACCCACCCC	3354
Qy	1041	ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg	1060
Db	3355	ACCATCTGTGTTCGGGGCATCTCTGTGGCGTGCATCTTCTCTCTGTGTGGGTTCG	3414
Qy	1061	ValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys	1080
Db	3415	ATGATCCGTTTACCAACCGCGTCACTGGCCCGGATCATCGAACGTGTGGCACTAATGC	3474

RESULT 5

```

US-10-209-059-21
; Sequence 21, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIORITY APPLICATION NUMBER: 60/096,822
; PRIORITY FILING DATE: 1998-08-17
; PRIORITY APPLICATION NUMBER: 09/371,383
; PRIORITY FILING DATE: 1999-08-06
; PRIORITY APPLICATION NUMBER: 09/550,483
; PRIORITY FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 21

```



```
QY 641 AlailleTyrglyTyrGluProProilleLysAlaLysLysProglyPheileuAlaSerLeu 660
Db 2158 GCTCTATATGGTTATGACCCCAATTAAGCAAAAGAGGGTGGTTCTTGTTCATCACTA 2217
QY 661 CysglyGlyLysLysLysAlaSerLysLysLysLysArgSerSerAspLysLysLysSer 680
Db 2218 TGTGGTGGCAGGAGGAGCAATCAAGAAG--GGCTCAGACAAGAAAAGTCA 2274
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2275 CAGAGCATGTGGACAGTCTCTGTGCCAGTATTCAATCTTGAAGATATAGAGGAGGAGTT 2334
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2335 GAAGCGCTGGATTGATGAGAAATCACTTCTTATGTCTCAATGAGCTTGAGAG 2394
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2395 AGATTGGCCAATCTGCAGCTTTTGTGCTCCACTCTGATGGAATATGGTGTCTCT 2454
QY 741 GlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2455 CAGCTCTGCAGCTCCAGATCTCTCTGAAGAAGAGCTATCCATGTTCATAAGTTGTGGCTAC 2514
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db 2515 GAGGACAGATTGAATGGGAACCTGAGATTGGGTGGATCTATGGTCTGTGACGGAAGAT 2574
QY 781 IleLeuThrGlyPheLysMethHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2575 ATTCTCACTGGGTTCAAGATGCACGACGAGGCTGGCGTGCATCTACTGCATGCCTAAG 2634
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2635 CGCGCGGCTCTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACACAGTGTCTC 2694
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2695 CGGTGGGCTCTCGTTCAGTGAATCTTTTCAGCGGCAATGGCCCTATGGTACGGG 2754
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2755 TACGAGGACGCGCTCAAGTTCTTGGAGAGATTGGCTACATCAACACCAACATCTACCCG 2814
QY 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2815 CTCAGTCCCTCCCGCTCCTCATTTACTGTATCCTCGCTGCCATCTGCTGTCTCAGCGG 2874
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2875 AAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGTATCTGGTTTCATCTCTCTTCATC 2934
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 2935 TCGATCTTCGCACCGGATCTCTGAGATGAGTGGAGCGGCGTGGGCATCGACAGTGG 2994
QY 921 TrpArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2995 TGGAGGAACGACGAGTTCCTGGGTCTATCGAGGCACTCTCGGCCACTCTTCGCCGTCTTC 3054
QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3055 CAGGCGCTCTCAAGGTCTTGGCGCATCGACCACTTCACCGTCCACCTCCAAGGCC 3114
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuLe 980
Db 3115 TCGATGAAGACGGGCACTTCGCGAGGTGTACATGTTCAAGTGACGACACACTTCTGTATC 3174
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3175 CGGCCCAACCACTCATCATCAACCTGTGGCGGTGTGTGGCGCATCTCTCCACGCC 3234
```

```
QY 1001 IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyr 1020
Db 3235 ATCACACGGGGTACAGTCTGTGGGGTCCGCTCTTCGGCAAGCTCTTCTTCGCTCTCG 3294
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3295 GTGATCGTTCACCTGTACCGTTCCTCAAGGGTCTCATGGGTCTGGCAGAAACCGCACCCG 3354
QY 1041 ThrIleValIleValTyrAlaValLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Db 3355 ACCATCGGGTGTCTGGCGCATCTGTGGCGTGCATCTTCTCTGCTGGGTTCGC 3414
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3415 ATCGATCCGTTCAACCAACCGGTCTACTGGCCCGGATATCTCGAACCTGTGGCATCACTGC 3474

RESULT 6
US-10-160-719-29
; Sequence 29, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3551)
US-10-160-719-29

Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 15 Gaps: 2

US-09-900-237A-30 (1-1080) x US-10-160-719-29 (1-3746)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 321 ATGGAGGCGCAGCGCGGCGTGAAGTCGGGAGGCGCGTGGCGGACAGGTGTGCCAG 380
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 381 ATCTCGCGCAGCGCGTGGCGCACCGCGGAGGGGAGCGTCTTCGCCGCTCGCACGTC 440
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 441 TCGGGTTCGCTGTGCGCCCTCTGCTACAGTACGAGCGCGGAGCGGACGCGACGCGCGG 500
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 501 TGCCCCCAGTGCAGACCAAGTACAAAGCGCCACAGGGGAGCGCCGCGCATCGTGGGGAG 560
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
```

Db 561 GAAGGAGCAGCACTGATGCCGAT-----AGCGACTTCAATTACCTTGTCATCTGGCAAT 614
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 615 GAGGACCAAGAGAGAGATGCGCAGAGAAATGCGCAGCTGGCGCATGAACCTTTGGGGGC 674
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 675 AGCGGGAGATTTGGTCGCCCCAGATGATGAGTGGCGAGATCGGGCTTTCACAGATGAC 734
QY 141 SerGlyGluLeuProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 735 AGTGGCGAGATTCCTCGGGGATATCCCATCAGTCACTAACACCGCAGATCTCAGAGAA 794
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 795 ATCCCTGGTCTCCCTGACCATCATATGATGTCGCCAACTGGGAACATTTGGCAACGT 854
QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 855 GCTCCATTTCCCTATGTGAACCATTCGCCAAATCCGTCAGGGAGTTCTCTGGTAGCAT 914
QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220
Db 915 GCGAATGTTCCCTGGAAAGAGAGGTTGATGGCTGGAAATGAAGCAGGACAAAGGGGACG 974
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 975 ATTCCCATGCAAGTGGCAACAGCATTTGCTCCCTCTGAGGGTTCGGGTGTTGGTAGATT 1034
QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 1035 GATGCATCACTGATACACATGGAAGATGCCCTATTGACGCGAATCTCGACACCT 1094
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
Db 1095 CTATCTAGGAAAGTTCACCTTCTCTCCAGGATAAATCCATACAGATGGTCAATTGTG 1154
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
Db 1155 CTGCGATGATGTTCTTAAGCATCTCTTGACATACCGTATCAAAATCTCTGCGCAAT 1214
QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
Db 1215 GCATACCATATAGGCTTCTATCTGTATATGTGAGATCTGGTTGCTCTTCTGTTGATA 1274
QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
Db 1275 TTGGATCAGTTCCTTAAGTGGTTTCCAAATCAACCGGGAGAGCTACCTTGATAGGCTGCA 1334
QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaValAspIlePheValSer 360
Db 1335 TTAAGGTATGACCGGAAGTGAGCCATCTCAGTTGGCTGCTGTGTGACATTTTGGTCAGT 1394
QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
Db 1395 ACAGTCACCAATGAAGAGACCTCTCTGTGCTACTGCCAATACCGTCTATCCATCTT 1454
QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
Db 1455 GCTGTGGATTACCTGTGGATAAGTCTCTTGTCTATGTATCTGATGATGAGCTGGCATG 1514
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
Db 1515 CTGACATTTGATGCACTAGCTGAGACTTCAGATTTGCTAGAAATGGGTACCATTTGTT 1574
QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
Db 1575 AAGAAGTACACATTTGAACCTTAGAGCTCTCGAATGGTACTTCTCCCGAATAATTGATTAC 1634
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
Db 1635 TTGAAGGACAAAGTGACCCCTTCATTTGTTAAAGACCGCGGGCCATGAAGAGATAT 1694

QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
Db 1695 GAAGAATTCAAAGTTAGGTAAATGGCTTGTGCTAAAGGCACAGAAAGTTCTCTGAGAA 1754
QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
Db 1755 GGATGGATCATGCAAGATGGCACACCATGGCCAGGAACAATACCAGGGACCATCTCTGGA 1814
QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
Db 1815 ATGATTCAGGTTTTCTTGGTTCACAGTGGTGGCTTGATCTAGGCAATAGCTACCC 1874
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
Db 1875 CGTTGGTCTATGTTTCTCGTGAAGCGTCTCGGATTCAGCATCAACAAGAAAGCTGGT 1934
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
Db 1935 GCCATGAATGCTCTTGTTCGTGCTCAGCTGTCTTACCAATGGACAATACATGTTGAAT 1994
QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db 1995 CTTGATTTGTCATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCTCT 2054
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db 2055 ATGGACCTTAACCTAGGAAGAGGTGCTGCTAGTCCAGTTTCCCGAGATTCGATGGC 2114
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db 2115 ATTGACAGGAATGATCGATATGCCACAGAAACACCGTGTTCGATATTAATTGAGTA 2174
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 2175 GGTCTTGATGGCATCCAAGGACAGATTATGTCCGAACCTGGCTGTGTTTCAACCGAACA 2234
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2235 GCTCTATGTTGTTATGAGCCCCCAATTAAGCAAGAAGGGTGGTTTCTTGTCTATCTACTA 2294
QY 661 CysGlyLysLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db 2295 TGTGGCGGTAGGAAGAGGCAACCAATCAAGAGAG---GGCTCGACACAGAGAAGTCG 2351
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2352 CAGAAGCATGTGGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTT 2411
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2412 GAAGGGCGCTGGATTTGACGACGAGAAATCACTTCTTATGTCTCAATAGCGCTGGAGAG 2471
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
Db 2472 AGATTTGGCAGTCCGACAGGTTTGTTCCTCCACTCTGATGGATGATGGTGGTTCCT 2531
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2532 CAGTCCGCAACTCCGGAGTCTCTTCTGAAAGAAGCTATCCATGTTATAAGCTGTGGCTAT 2591
QY 761 GluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAsp 780
Db 2592 GAGGACAAGACTGAATGGGAACGTAGATCGGGTGGATCTACGGTTCTGTGACAGAGAGAC 2651
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLys 800
Db 2652 ATTCTACCGGATTCAGATGCGCGGAGGCTGGCGGTGATCTACTGATGCCACAG 2711
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2712 CGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTTCGACACCTCTGTAACAGGTGCTC 2771


```
Qy 821 ArgTrpAlaLeuGlySerValGluLeuPheSerArgHisCysProLeuTrpTyrGly 840
Db 2772 CGGTGGGTCTTGGTCCGTGGAGATCTCTTCACCGGCACTGCCCTCTGGTACGGC 2831
Qy 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2832 TACGAGGCGCGCTCAAGTTCCTGGAGAGATTGGGTATCATCAACACCACTCTACCG 2891
Qy 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2892 CTCAGTCTCATCCCTCTCTCATCTACTACTGATCTTCCCGCCCATCTGTCTGCTACCGGA 2951
Qy 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db 2952 AAGTTCATCATCCAGAGATCAGCAACTTCGCCAGCACTGGTTCATCTCCCTCTCATC 3011
Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
Db 3012 TCGATCTTTCGCCACCGGCATCTCTGGAGATGAGTGGAGCGGGTGGGCATCGACGAGTGG 3071
Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 3072 TGGAGAACGAGCAGTCTTGGGTGATCGGGGCACTCTCGGCCACTCTTTCGCCGTGTTC 3131
Qy 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3132 CAGGGCTTGTCTCAAGGTCTGGCCGGCATCGACACCACTTCACCGTCACTCCCAAGGCC 3191
Qy 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrLeuLeuIle 980
Db 3192 TCGGACGAGGACGGGCACTTCGCGAGGTGTACATGTTCAAGTGACGACGCTCTGTATC 3251
Qy 981 ProProThrThrIleLeuIleIleAsnMetValGlyValAlaValAlaGlyThrSerTyrAla 1000
Db 3252 CGGCCACACCATCTGATCATCACTGATGCTGGTGGCGTCTCGCCGGCATCTCTTACGCC 3311
Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db 3312 ATCAACAGCGGATACAGTCTGGTGGGCGCGCTCTTCGGCAAGCTCTTCTTGGCCCTTGG 3371
Qy 1021 ValIleValHisLeuTyrProPheLeuLysGlyIleLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3372 GTCATGCTGCACCTGATCCGCTTCTCAAGGCGCTCATGGGAGGAGCAACCGCACCGCG 3431
Qy 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3432 ACCATGCTGCTCTGGGCCATCTGCTGGCTCCATCTTCTCTGCTGTGGGTTCGC 3491
Qy 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3492 ATCGACCCCTTCACCAACCGCGTCACTGGCCCGGATACCCAGACGCTGTGGCATCAACTGC 3551
```

RESULT 7

```
US-10-160-719-49
; Sequence 49, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
```

```
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719-49
Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservatives: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 15 Gaps: 2
US-09-900-237A-30 (1-1080) x US-10-160-719-49 (1-3746)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 321 ATGGAGGCGCAGCGCGACGGGTGAAGTCGGGAGCGCGGTGGCGGACAGGTGTGCCAG 380
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysaspVal 40
Db 381 ATCTCGCGCGCAGCGGTGGGCACCAACGCGCGGAGGGGAGCGTCTTCGCCCTCGCACGTC 440
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 441 TCGCGGTTCCTGGTGTGCGCCCTCTGCTACGAGTACGAGCGGACGAGCGGACGCGGCG 500
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 501 TGCCCCCAGTGCACAGCAAGTACAAAGCGCCACAGGGGAGCGCGCGCATCGTGGGAG 560
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 561 GAAGAGAGCAGCACTGATGCGCAT-----AGCGACTTCAATTACCTTGCATCTCGCAAT 614
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 615 GAGGACCAAGAGAGAGAGATTGCCACAGATTGCCAGCTGGCGCATGAACTGGGGGC 674
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 675 AGCGGGGATGTTGGTCCGCCCAAGTATCACAGTGGCGAGATCGGCTTACCAAGTATGAC 734
Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 735 AGTGGCGAGATTCTCTGGGGATACATCCCATCAGTCACTAAACAGCCAGATCTCAGGAGAA 794
Qy 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 795 ATCCCTGGTGTCTTCCCTGACCATCATATGATGTCCTCCCACTGGGAAACATTGGCAAGCT 854
Qy 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 855 GCTCCATTTCCCTATGTGAACCAATTCGCCAAATCCGTCAAGGAGTCTCTGTGTAGCATT 914
Qy 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGluAspLysGlyAla 220
Db 915 GGGATGTTGCTCGAAAGAGAGGGTGTGCTGGTGAATAATGAACAGGAGCAAGGGGAGC 974
Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 975 ATTCCTATGACGAATGGCACAAGCATCTCTCTCTGAGGGTTCGGGTGTGGTATATT 1034
Qy 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 1035 GATGATCAACTGATTACAAATGAGATGCTTTATTGAACAGCAAGAACTGACAGCT 1094
Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
Db 1095 CTATCTAGGAAGTTCCCACTCTCTCTCCAGATAAATCCATACAGGATGTGTCATTGTG 1154
```



```
Db 3312 ATCAACAGCGGATACAGTCGTGGGGCGCGCTCTTCGGCAAGCTCTTCTTGCCCTTCTGG 3371
Qy 1021 VallileValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3372 GTCAFCGTCCACCTGTACCGCTTCCTCAAGGCGCTCATGGGCGAGCAAGACCGACCCCG 3431
Qy 1041 ThrileValIleValTriaValleuLeuAlaSerilePheSerLeuLeuTyrValArg 1060
Db 3432 ACCATCGTGTGTCTGGGCCATCTCTGGGGTCCATCTTCTCTGCTGTGGGTTCGC 3491
Qy 1061 VallAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3492 ATCGGCCCTTCACCAACCGCGTCACTGGCCCGGATACCCAGACGTGTGGCATCAACTGC 3551

RESULT 8
US-10-627-132-9
; Sequence 9, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Haiyin
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; PRIORITY FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-9

Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 13 Gaps: 2

US-09-900-237A-30 (1-1080) x US-10-627-132-9 (1-3773)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 338 ATGAGGCGGACGCGGACGGCGTGAAGTCGGGGAGCGCGTGGCGGACAGGTGTGCCAG 397
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 398 ATCTCGCGGACGGCGTGGGACCAACCGCGAGGGGAGCGTCTTCGCCGCTTCGACGCTC 457
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 458 TCGGGGTTTCGGGTGTGCGCCCTGCTGCTAGGTAGTAGGACGACGAGCGGACGCGAGCG 517
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 518 TGCCCCCAGTGCAGACCAAGTACAAAGCGCCACAAAGGGGAGCGCCGCGATCGGTGGGAG 577
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 578 GAAGGAGACGACACTGATGCCGAT-----AGCGACTTCAATTACTTGTGCATCTGCCAAT 631
```

```
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 632 GAGGACACAGAGCAGAGCAAGATTGCCACAGAAATCGCAGCTGGCGCATGAACGTTGGGGC 691
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 692 AGCGGGGATGTGTGTCGCCCAAGTATGACAGTGGCAGATCGGGCTTACCAGATATGAC 751
Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 752 AGTGGCGAGATTCTCGGGGATACATCCCATCAGTCACTAACAGCCAGCATCTCAGAGAA 811
Qy 161 IleProGlyAlaSerProAspHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 812 ATCCCTGGTGTCTTCCCTGACCATCATATGATGTCCTCCCACTGGGAACATTGGCAAGCGT 871
Qy 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerile 200
Db 872 GCTCCATTTCCCTATGTGAACCATTCGGCAAAATCCGTCAAGGAGATTCTCTGTGTAGCAT 931
Qy 201 GlyAsnValAlaTatLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
Db 932 GGGAAATGTGCTGGAAAGAGAGGTTGATGCTCGAAATGAACAGACAGCAAGGGAGC 991
Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 992 ATTCCTCATGACGATGGCACAAGCATTCCTCCCTGAGGGTGGGGTGTGGTGATATT 1051
Qy 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 1052 GATGCATCAACTGATTACAAATGGAAGATGCTTTATTGAACAGCAAGCAACTCGACAGCT 1111
Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
Db 1112 CTATCTAGAAAGTTCCACTCTCTCTCCAGGATAAATCCATACAGATGGTCAATTG 1171
Qy 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
Db 1172 CTGCGATTGATTGTTCTAAGCATCTTCTTGCACTACCGTATCACAAATCCTCTGCGCAAT 1231
Qy 301 AlaTyrProLeuTyrPheLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle 320
Db 1232 GCATACCCATTTATGGCTTCTATCTGTATATGTGAGATCTGGTTCCTTCCTCGTGATA 1291
Qy 321 LeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
Db 1292 TTGGATCAGTTCCTCCCTAAGTGGTTTCCAAATCAACCGGGAGACGTACTCTGTAGAGTGGCA 1351
Qy 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
Db 1352 TTAAGGTATGACCGGGAGGTTGAGCATCTCAGTTGGCTGCTGTGTGACATTTTCGTCAGT 1411
Qy 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
Db 1412 ACAGTCGACCAATGAAGGAGCCTCTTGTGCTACCTGCAATACCGTCTATFCAATCTT 1471
Qy 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
Db 1472 GCTGTGGATTACCTCTGGATGAAGTCTCTTCTGTAAGTATCTGATGATGAGCTGGAGT 1531
Qy 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
Db 1532 CTGACATTTGATGCATAGCTGAGACTTCAGACTTTGTAGAAAATGGGTACCAATTTGTT 1591
Qy 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
Db 1592 AAGAGTACACATTTGAACCTTAGAGTCTCTGTAAGTCTCTTCTCCCAAAAATGATTAC 1651
Qy 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
Db 1652 TTGAGGACAAAGTGCACCTTTCATTGTTAAAGCCCGCGGCCCATGAAGAGAAATAT 1711
Qy 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
```

```
Db      1712 GAAGATTCAAAGTTAGGTAAGTGGCTTGTGCTTAAGCAGACAGAAAGTTCTCTGAGAA 1771
Qy      481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
Db      1772 GGATGGATCATGCAAGATGACACACCATGCGCAGGAAACAATACCGAGGACCATCTCGGA 1831
Qy      501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
Db      1832 ATGATTGAGTTTCTTGTGTACAGTGGTGGCTTGTATCTAGGGCAATAGCTATCCC 1891
Qy      521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
Db      1892 CGTTTGGTCTATGTTCTCTGTAAGAGCGTCTGGATTCCAGCATCACAGAAGCTGGT 1951
Qy      541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
Db      1952 GCCATGAATCTCTTGTGTCTCAGCTGTGCTTACCAATGACAAATACATGTTGAAT 2011
Qy      561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db      2012 CTTGATTGTGATCACTACATTAACAAGTAAGCTCTCAGGAAGCTATGTGCTTCCTT 2071
Qy      581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db      2072 ATGGACCTTAACCTAGGAAGAGTGTCTGCTACGTCCAGTTTCCCCAGAGATTCTGATGC 2131
Qy      601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db      2132 ATTGACAGGAATGATCGATATGCAACAGGAACACACCGTGTGTTTTCGATATTAACCTGAGA 2191
Qy      621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db      2192 GGTCTTGATGGCATCCAAAGCAGCATTTATGTGGAACTGGCTGTGTTTCAACCCGAACA 2251
Qy      641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db      2252 GCTCTATATGTTATGAGCCCAATTAAGCAGAGAGGGTGGTTCTTGTTCATCACTA 2311
Qy      661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerAspLysLysLysSer 680
Db      2312 TGTGGCGGTAGGAAGAGGCAAGCAATCAAGAAG---GGCTCGGACAAAGAAGATCG 2368
Qy      681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db      2369 CAGAAGCATGTGGACAGTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTT 2428
Qy      701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db      2429 GAAGGGCTGGATTTGACGACGAGAAATCACTTCTTATGTCCTCAATGAGCCTGGAGAAG 2488
Qy      721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db      2489 AGATTTGGCCAGTCCGACAGCTTGTGTGCTCCACTCTGATGGAGTATGTGTGTTCCT 2548
Qy      741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db      2549 CAGTCCGCACATCGGAGTCTCTTCTGAAGAAGCATATCCATGTTATAAGCTGTGGCTAT 2608
Qy      761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db      2609 GAGGACAAGACTGATGGGNACTGAGATCGGTGGATCTACGGTCTGTGACAGAGAC 2668
Qy      781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLys 800
Db      2669 ATTCACCGGATTCAGAGATGCAACGCGAGGCTGGCGGTGATCTACTGATGCCCAAG 2728
Qy      801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db      2729 CGGCCAGCTTTCAGGGGTCTGCCCCCAATCAATTTTCGAGCCGCTCTGAACCCAGGTCTC 2788
Qy      821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrGly 840
```

```
Db      2789 CGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCGCGCACTGCCCTGTGGTACGGC 2848
Qy      841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db      2849 TACGGAGGGCGGTCTAAGTTCTCTGGAGAGATTCGGGTACATCAACACCATCTATCCCG 2908
Qy      861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db      2909 CTCAGTCCATCCGCTTCTCATCTACTGATCTCTGCCGCCATCTGTCTGTCTACCGGA 2968
Qy      881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db      2969 AAGTTTCATCATCCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3028
Qy      901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTyr 920
Db      3029 TCGATCTTCGCCACGGGCATCTCTGGAGATAGGTGGAGCGGGGTGGGCATCGCAGAGTGG 3088
Qy      921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
Db      3089 TGGAGGNAACGAGCAGTTCTGGGTGATCGGGGGCATCTCCGGGCACCTCTTCGCGGTTC 3148
Qy      941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db      3149 CAGGGCTGCTCAAGGTGCTGGCGGCATCGACACCAACTTCACCGTCACCTCCAGAGCC 3208
Qy      961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db      3209 TCGAGCAGAGACGGCGACTTCGCGGAGCTGTACATGTTCAAGTGGAGCGCTCTCTGATC 3268
Qy      981 ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db      3269 CCGCCACACCATCTGATCATCACTGTCGGGTGTCGCGCGCATCTCTCTAGGCC 3328
Qy      1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrp 1020
Db      3329 ATCAACAGCGGATACAGTGTGGGGCCGCTCTTGGCAAGCTCTTCTTTCGCTCTCG 3388
Qy      1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db      3389 GTCATCGTCACCTGATCCCGTTCCTCAAGGGCTCATGGGCGCATCAAGACCGCACCCCG 3448
Qy      1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db      3449 ACCATGTCGTCTCTGGGCCATCTCTGTCGGCTGCATCTTCTCTTCTGTGGTTCGC 3508
Qy      1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db      3509 ATCGACCCCTTCACCCCGCTCACTGGCGCGGATACCCAGACGTGTGGCATCAACTGC 3568
```

RESULT 9

```
US-10-209-059-9
; Sequence 9, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Haiyin
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
```


Db 2312 TGTGGCGTAGGAGAGAGGCAAGCAATCAAGAG--GGCTCGGACAGAGAGAGTGG 2368
QY 681 AsnLysHisValAspSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2369 CAGAAGCATGTGCACAGTTCGTGTGCCAGTATTCAACCTTGAGAGATATAGAGGAGGAGTT 2428
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2429 GAAGGGCGCTGGATTGACGACGAGAAATCACTCTTATGTCTCAAAATGAGCCCTGGAGAG 2488
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2489 AGATTGGCCAGTCCGAGGCTTGTGTCTCACTCTGATGGAGTATGGTGTCTCT 2548
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2549 CAGTCCGCAACTCCGGAGTCTCTTCTGAAAGAGCTATCCATGTTATAAGCTGTGGCTAT 2608
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrPheIleTyrGlySerValThrGluAsp 780
Db 2609 GAGGACAAGACTGAATGGGAACTGAGATCGGGTGGATCTACGGTCTGTGACAGAAAGC 2668
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2669 ATTCTACCGGATTCAGATGCAGCGCGAGGCTGGCGGTGATCTACTGATGCCCAAG 2728
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2729 CGGCCAGCTTCAAGGGGTCTGCCCCCATCAATCTTTCGAGCCGTCTGAACCAAGTCTC 2788
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2789 CGGTGGGCTCTTGGGTCCGGAGATCCTCTTCAGCCGGCACTGCCCCCTGTGGTAGCGC 2848
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2849 TAGCGAGGCGGCTCAAGTTCCTGGAGAGATTCGGTACATCAACACCAACCATACCCG 2908
QY 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2909 CTCAGTCCATCCGCTTCTCATCTACTGATCTCGCCGCCATCTCTCTCTCACCGGA 2968
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2969 AAGTTCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3028
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 3029 TCGATCTTCGCCAGCGCATCTCGAGATGAGTGGAGCGGGTGGGCATCGACGAGTGG 3088
QY 921 TyrArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 3089 TGGAGGAACAGAGAGTTCGGGTGATCGGGGATCTCCGCGCACTCTTCGCGCTGTC 3148
QY 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3149 CAGGCTCTCTCAAGGTGCTGGCGGATCGACACCACTTCACCGTCACCTCCAGGCC 3208
QY 961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle 980
Db 3209 TCGGACGAGGAGCGGACTTCGCGAGCTGTACATGTTCAAGTGGACGAGCTCTCTATC 3268
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3269 CCGGCCACCACTCTGATCATCACTGTGCGCGCTCGTCGCGGCGCATCTCTACGCC 3328
QY 1001 IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPheAlaPheTyr 1020
Db 3329 ATCAACAGCGGATACCAAGTGTGGGGCCGCTCTTCGCGCAAGCTCTCTTCGCTCTCG 3388
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040

Db 3389 GTCATCGTCCACCTGTACTCCGTTCTCAAGGGGCTCATGGGAGAGAACCGCACCCCG 3448
QY 1041 ThrIleValIleValTyrAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Db 3449 ACCATCGTCTGCTCTGGGCGCATCTGCTGGCGTCCATCTCTCTTGTGTGGTTCGC 3508
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAspCys 1080
Db 3509 ATCGACCCCTTACCACCCCGGTACTGGCCCGGATACCCAGAGCTGTGGCATCACTGC 3568
RESULT 10
US-10-160-719-5
; Sequence 5, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338) ... (3568)
US-10-160-719-5

Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 15 Gaps: 2

US-09-900-237A-30 (1-1080) x US-10-160-719-5 (1-3773)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 338 ATGAGGGCGACCGGACCGGCGTGAAGTTCGGGGAGCGCGGTGGCGACAGGTGTGCCAG 397
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 398 ATCTGCGCGACGCGGTGGGCACCAACCGCGGAGGGGACGCTCTTCGCGCTGCGAGCTC 457
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 458 TGGCGGTTTCGGGTGTCCGCCCTGCTACGAGTACGAGCGCAAGGACGCGCAGCGCG 517
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 518 TGCCCCAGTGCAGACCAAGTCAAGCGCCCAAGGGGAGCCCGCGCATCCCGTGGGGAG 577
QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 578 GAAGGACGACACTGATGCCGAT-----AGCGACTTCAATTACCTTGATCTGGCAAT 631
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db 632 GAGGACCAAGCAGAAAGATTGCCGACAGAAATGCCGAGCTGGCGCATGAACGTTGGGGGC 691

Db 2849 TACGAGGGCGGCTCAAGTTCTCGGAGAGATTCCGGTACATCAACACCACTATCCCG 2908
Qy 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2909 CTCACGTCATCCGCTCTCATCTACTGATCTCGCCGCCATCTGTCTGTCTCACCGGA 2968
Qy 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleThrPheIleAlaLeuPheLeu 900
Db 2969 AAGTTTCATCATCTCCAGAGATCAGCACTTCGCCAGCATCTGGTTCATCTCCCTCTTC 3028
Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
Db 3029 TCGATCTTCGCCAGGCGATCTCGAGATGAGGTGGAGCGGGTGGGCGATCGACAGTGG 3088
Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaIleSerAlaIleValPhe 940
Db 3089 TGGAGGAACAGCAGATTCTGGGTGATCGGGGCGATCTCCGCGCATCTCCGCTCTTC 3148
Qy 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3149 CAGGGCTGTCTCAAGGTGCTGGCGGCGATCGACACCACTTCACCGTCACCTCCAGGCC 3208
Qy 961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db 3209 TCGACGAGACGCGGACTTCGCGGAGCTGTACATGTTCAAGTGGAGCGCTCTCTGATC 3268
Qy 981 ProProThrThrLeuIleLeuAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3269 CCGCCACCACTCTGATCATCAACTGTGTGGCGCTGTGTGGCGGCGATCTCTCTAGGCC 3328
Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db 3329 ATCAACAGCGGATACCACTGTGTGGGCGCGCTCTTCGCGAAGCTCTTCTTCGCTCTCG 3388
Qy 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3389 GTCATCTGCACCTGTATCCCGCTTCTCAAGGCGCTCATGGGCGAGACCGCACCCCG 3448
Qy 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3449 ACCATCTGCTGCTGGGCGATCTCTGCGCGTCCATCTTCTCTGCTGTGGGTTCGC 3508
Qy 1061 ValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3509 ATGACCCCTTCACACCGCGTCACTGGCCCGGATACCCAGACGTGGGATCAACTGC 3568

RESULT 11

US-10-627-132-13
; Sequence 13, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-13

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 5149.00 Matches: 948
Percent Similarity: 94.08% Conservatives: 69
Best Local Similarity: 87.70% Mismatches: 58
Query Match: 89.11% Indels: 6
DB: 13 Gaps: 6

US-09-900-237A-30 (1-1080) x US-10-627-132-13 (1-3704)

Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 272 ATGACGGC---GGCGAGCGCCAGCAATTCGGGGAAGCATGTGGCGGAGGTGTGCAG 328
Qy 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 329 ATCTGGCGGACGCGGTGGGCGACCGCGCGCGGCGACCTCTTACCGCTGCGAGGTC 388
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
Db 389 TGGGCTTCCCGTGTGCGCCCATGCTACGAGTACGAGCGCAAGGACCGCACCGCGC 448
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 449 TGCCCGCAGTGCAGACTAAGTACAGCGCCACAAAGGAGCCACCAGTACACGGTGAG 508
Qy 81 GluGlyAspAspThrAspAlaAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 509 GMAATGAGCATGTGGATGTGACATGTGAGTGACTACAACTACCAAGCATCTGGCAAC 568
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 569 CAGGATCAGAAAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGACAACTACGTTGC 628
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 629 AGT---GATATTGGCTGGCTAAGTATGACAGCGGTGAATTTGGCGCATGGGAGTATGAC 685
Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 686 AGTGTGAGATCCCTGCTGTGATATCCCTCACTAACTCATAGCCAGATCTCAGAGAGAG 745
Qy 161 IleProGlyAlaSerProAsnHisHisMetSerProThrGlyAsnIleSerArgArg 180
Db 746 ATTCTCGAGCTTCCCTCTGAT---CATATGATGCTCTTGTGGGAAATTTGGCAGGCGT 802
Qy 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 803 GGACATCAATTTCTTATGTAATCATTTCCAAACCCATCGAGGGAGTTCTCCGGTAGC 862
Qy 200 IleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGly 219
Db 863 CTTGGCAATTTGTCATGAAAGAGAGGAGGTGGATGGATGGAATAATGAAG--GATAAAGGT 919
Qy 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
Db 920 GCAATTCCTATGACCAATGGAAAGCAAGCATTTGCTCCATCAGAAGGCGGTGGAGTTGCTGAT 979
Qy 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 980 ATTGATGCTTCTACTGATTATACATGGAAGATGCTTACTGATGATGAACCTCGGCA 1039
Qy 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
Db 1040 CCTCTACTAGAAAGTGCCAAATTCCTTCATCCAGATAAATCCGTACAGAAATGGTCAATT 1099
Qy 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
Db 1100 GTGCTAGCTTGGCTGTCTTCTATGATCATATCTTGGCTACCGGTATCACACATCTCTGTAAC 1159
Qy 300 AsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrp 319

Db 1160 AATGCATATCCACTGTGGCTTTTATCCGTCATATGTGAGATCTGGTTGCTTTGTCTCTGG 1219
Qy 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTrpLeuAspArgLeu 339
Db 1220 ATTTTGGATCAGTTTCCCAAGTGGTCCCAATCAACCGTGAACAATACCTTGTATAGACTG 1279
Qy 340 AlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheVal 359
Db 1280 GCTTTAAGGTATGACCGAAGGTGAACCATCTCAATTAGTCTCTGTGATATTTTGTGTC 1339
Qy 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
Db 1340 AGTACTGTGGATCCCAATGAAGAGGCTCTCTGTGCTACTGCAAACTACTGTGCTTTCCATC 1399
Qy 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
Db 1400 CTTGCTGTGCAATATCCGGTTGACAGGATCTTGTCTATGTTTCGATATGAGGTGCT 1459
Qy 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPhe 419
Db 1460 ATGCTGACTTTTGATGCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGGTTCGGTTC 1519
Qy 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
Db 1520 TGTAAAGAAAGTACAAATAGAGCTTAGGCCCGCGAATGGTACTTTGCTCAGAAAATTGAT 1579
Qy 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu 459
Db 1580 TACTTGAAGACAAAGTTCAAACCTCAATTTGTGAAGAAGACGCGGCGCAATGAAGAGANA 1639
Qy 460 TyrGluGluPheLysIleAArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
Db 1640 TATGAAGAATTCAAAGTTTCGTATCAATGCTCTGTAGCCAGGCACAAAAGTTCCCGAG 1699
Qy 480 GluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 499
Db 1700 GAGGATGGATCATCAGATGGTACACCTGGCCCTGGGAAACAATACTAGGGACCATCT 1759
Qy 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
Db 1760 GGAATGATTCAGTTTTCCTGGTACAGTGGAGGGCTTGCGTTGAAGGCATGAACAT 1819
Qy 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 539
Db 1820 CCTCGTTGGTTTATGTGCTCGTCAAAAACGCTCTGGATTCCAACATCAACAAGAGCT 1879
Qy 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
Db 1880 GGTGCCATGAATGCATCTGCTGTATCAGCTGCTCTTACTAATGGGCAATACATGTTG 1939
Qy 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
Db 1940 AATCTTGATGTGACCACTACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTGCTTC 1999
Qy 580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
Db 2000 CTTATGGACCCAAACCTAGGAAGGAATGCTGTTATGTCCAATTTCTCCAGAGGTTTGTAT 2059
Qy 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPheAspIleAsnLeu 619
Db 2060 GGTATGTATGAAGATGACCGATATGCAAAACAGGAACACTGTGTTTTCGATATTAACCTG 2119
Qy 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 2120 AGAGTCTTGACGCATTCAGGGCCAGTTTATGTGGAACTGGTTGTGTGTATTAACAGA 2179
Qy 640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2180 ACGGCTTATATGGTTATGAGCCCTCAGTCAAGAAAAAAGCCAGGCTTCTTCTCTTCG 2239
Qy 660 LeuCysGlyGlyLysLysAlaSerLysSerLysLysArgSerAspLysLysLys 679
Db 2240 CTTTGTGGGGAGGAGAAAAGACGTCAAAATCTAAGAAAG---AGCTCGAAGAAAGAGAG 2296

Qy 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGly 699
Db 2297 TCACATAGACACGACAGACTTCTGTACCAAGTATTTAATCTCGAGATATAGAGGAAGG 2356
Qy 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2357 ATTGAAGGTTCTCAGTTTGATGATGAGAAATCGTGATATGCTCAATAGACTGGAG 2416
Qy 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2417 AAGAGATTGGCCAGTCCAGTGTGTTTGTAGCTCTACTCTGATGGAATATGTTGTGT 2476
Qy 740 ProGlnSerSerThrProGlnSerLeuLysGluAlaIleHisValLysSerCysGly 759
Db 2477 CCACAACTGCAACTCCAGAGTCTCTTCTGAAGAAGACTATTCATGTCATCAGCTGTGGC 2536
Qy 760 TyrGluAspLysSerGluTrpGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db 2537 TATGAGGACAAAACCTGACTGGGAACTGAGATTGGGTGGATCTATGTTCTGTTCAGAA 2596
Qy 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2597 GACATTCTCACCGGATTCAGATGTCATGCTCGAGCTCGCGATCAATCTACTGCATGCT 2656
Qy 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2657 AAGCGACAGACTTTCAGGGATCTCTCTCTATCAACCTTCGATCGTGTGTAATCAAGTG 2716
Qy 820 LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr 839
Db 2717 CTTCCGTGGGCTCTTGGTTCATTTGAAATTTCTTTTCAGCAGGCAATGTCATATGTTAT 2776
Qy 840 GlyTyrGlyGlyArgLeuLysPheLeuArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2777 GGTATGAGGCGCGCTTAAATTCCTGAGAGATTGCTTATATCAACACAAATTTAT 2836
Qy 860 ProLeuThrSerProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2837 CCACCTCATCAATCCCGCTCTCTGTACTGTCATATTCGACAGAGTTGTCTTCTCACT 2896
Qy 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe 899
Db 2897 GGGAGATTTCATCATCCCAAGATTAGTAACCTAGAGAGTTGGTTGTTATATCGCTCTT 2956
Qy 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
Db 2957 ATCTCAATCTTTGCCACTGGTATCTTCAGATGAGGTGGGTGGTATTCATGAA 3016
Qy 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaVal 939
Db 3017 TGGTGGAGAACGACGACTTCTGGGTCAATGGTGGTATTTCTGCGCATTTATTTGCCGTC 3076
Qy 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3077 TTCAGGGTCTCTGAAGTCTGTTGCTGATCGACACAGAGCTTCACTGTCACTCTAAG 3136
Qy 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrLeuLeu 979
Db 3137 GCCACTGACGAAGAAGGTGATTTTTCGCGAGCTCTACATGTTCAAGTGGACAAACGCTTCG 3196
Qy 980 IleProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3197 ATCCCAACCAACTATTTTTCATCATCACTGCTGGTGGGCTGGTGGCTGGCATTTCTTAC 3256
Qy 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3257 GCAATCAATACGGGTTACAGTCATGGGACCTCTTTTCGGGAAAGCTCTTCTTTCGCTC 3316
Qy 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGluAsnArgThr 1039
Db 3317 TGGTGTATTGTCACCTGTATCCCTTCTCAAGGCGCTCTAAGGGGACAGACAAACCGCACG 3376


```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)....(3499)
US-10-160-719-57

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 5149.00 Matches: 948
Percent Similarity: 94.08% Conservative: 69
Best Local Similarity: 87.70% Mismatches: 58
Query Match: 89.11% Indels: 6
DB: 15 Gaps: 6

US-09-900-237A-30 (1-1080) x US-10-160-719-57 (1-3704)

QY 1 MetAspGlyAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 272 ATGGACGGC---GGCGACGCCACGAATTCGGGAAGCATGTGGCGCGCAGGTGTGCCAG 328
QY 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 329 ATCTGGCGGCACGGCGTGGGCACCGCGCGGACGGCGACCTCTTCACCGCGCTGCGACGTC 388
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 389 TGGGGTTCCTCCGTGTGCCGCCCATGCTACGAGTACGAGGCGCAGGACGCCACCGAGCG 448
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 449 TGGCGCGAGTGCAGACTAAGTACAAAGCCCAACAAAGGGAGCCCAACAGTACACGGTGAG 508
QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 509 GAAATAGGAGTGGAGTGTGACGATGTGAGTGTGAGTACTACAACTACCAAGCATCTGGCAAC 568
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db 569 CAGGATCAGAAAGCAAGAGTGTCTGAGAGATGCTCACTTGGCGGACAACTACAGTGGC 628
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 629 AGT---GATATTGSCCTGGCTAAGTATGACAGCGGTGAATTTGGCGATGGGAAGTATGAC 685
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 686 AGTGGTGAGATCCCTCGTGTGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAG 745
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 746 ATTCTGGAGCTTCCCTGTAT---CATATGATGTCTCTGTGGGAACATTTGGCAGCGCT 802
QY 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 803 GGACATCAATTTCTTATGTAATCAATCTCCAAACCCATCGAGGGGAGTCTCCGGTAGC 862
QY 200 IleGlyAsnValAlaIleProLysGluArgValAspGlyTyrLysMetLysGlnAspLysGly 219
Db 863 CTTGGCAATGTTCATGCAAGAGAGGGTGGATGGATGGAAATGAAG---GATAAAGGT 919
QY 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
Db 920 GCATTCCTATGACCAATGACNAGCATTTGCTCCATCAGAGGGCGGTGGAGTTGCTGAT 979
QY 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 980 ATTGATGCTTCTACTGATTAATAACATGGAAGATCCCTTACTGATGATGAACCTCGCAA 1039
QY 260 ProLeuSerArgLysValProIleAlaSerSerLysLysIleAsnProTyrArgMetValIle 279
Db 1040 CCTCTATCTAGAAAGTGCCAAATTCCTTCATCCAGAAATAAATCCGTACAGAATGGTCAAT 1099
QY 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299

```

```
QY 660 LeuCysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLys 679
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2240 CTTTGTGGGAGAGAAAGACGTCATAATCTAAGAG--AGTCGGAAGAAAGAG 2296

QY 680 SerAsnLysHisValAspSerValProValPheAsnLeuGluAspIleGluGly 699
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2297 TCACATAGACACGCAGACAGTCTCTGACCAGTATTAAATCTCGAAGATATAGAGAGG 2356

QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2357 ATTGAAGGTTCTCAGTTTCATGATGAGAAATCGCTGATTATGTCCTCAATGAGCTTGGAG 2416

QY 720 LysArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2417 AAGAGATTGGCAGTCAGTGTCTTTGTAGGCTCTACTCTGATGGAATATGTTGGTGT 2476

QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2477 CCACAATCTGCAACTCCAGAGTCTCTTCTGAAAGAAGCTATTTCATGTCATCAGCTGTGC 2536

QY 760 TyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2537 TATGAGGACAAACTGACTGGGAACTGAGATGGGTGGATCTATGGTTCTGTTACAGAA 2596

QY 780 AspileLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2597 GACATCTCACCGGATTCAAGATGATGCTCGAGGCTGGCGATCACTACTGATGCT 2656

QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2657 AAGGACACAGCTTCAAGGATCTGCTCTATCAACCTTCGGATCGTTTGAATCAAGTG 2716

QY 820 LeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyr 839
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2717 CTTCCGTGGGCTCTTGGTTCATTGAAATCTTTTCAGCAGGCAATGTCCTATGGTAT 2776

QY 840 GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2777 GCGTATGAGGCGCGCTTAATCTCTGGAGAGATTGCTTATATCAACACAACTTAT 2836

QY 860 ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2837 CCACCTCACATCAATCCGCTCCTCCTCTACTGCATATGCCAGCAGTTTGTCTTCACT 2896

QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe 899
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2897 GGGAAAGTTTCATCATCCAAAGATTAGTAACCTAGAGAGTGTGTGTTTATATCGCTCTTT 2956

QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGlu 919
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
2957 ATCTCAATCTTGCCACTGGTATCTCTGAGATGAGGTGAGTGGTGGCATTCATGNA 3016

QY 920 TrpTyrArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3017 TGGTGGAGAACGACGAGTTCCTGGGTCTATTTGGTGTATTTCGCGCAATTTATTTCCGCTC 3076

QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3077 TTCAGGGTCTCCTGAAGGTGCTTGTGTGTATCGACACAGGCTTCACGTCACTCTAAG 3136

QY 960 AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeu 979
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3137 GCCACTGACGAAGAAGGTGATTTTCCGAGGCTCTACATGTTTCAAGTGGACACGCTTCTG 3196

QY 980 IleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3197 ATCCCAACAACCATATTGATATCATCAACCTGGTGGCGGTGGTGGCTGCTGCTAC 3256

QY 1000 AlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3257 GCAATCAATAGCGGTTACAGTCATGGGACCTCTTTTCGGGAAGCTCTCTTTTCGGTTC 3316
```

```
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3317 TGGGTGATTGTCACCTGTACCCCTTCTCAAGGGCTCATGGGAGACGACGACGACG 3376

QY 1040 ProThrIleValIleValTyrAlaValLeuAlaSerIlePheSerLeuLeuTyrVal 1059
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3377 CGACCAATTGTGTGTCTGGGCTATCTCTCTTGGCTGATCTTTTCCGTGATGTGGT 3436

QY 1060 ArgValAspPropheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3437 CGATTCGATCCATTCAACCCGGGTCACTGGCCCTGATATCGGAAATGTGGCATCAAC 3496

QY 1080 Cys 1080
Db |||
3497 TGC 3499

RESULT 14
US-10-425-114-2553
; Sequence 2553, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2553
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700217152_FLI
US-10-425-114-2553
```

```
Alignment Scores:
Pred. No.: 0 Length: 3534
Score: 5129.00 Matches: 944
Percent Similarity: 93.83% Conservative: 71
Best Local Similarity: 87.33% Mismatches: 60
Query Match: 88.77% Indels: 6
DB: 13 Gaps: 6
```

US-09-900-237A-30 (1-1080) x US-10-425-114-2553 (1-3534)

```
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 132 ATGGACGCGC---GGCGACGCCAGAAATTCGGGGAAGCATGTGCCGGGAGGTGGCCAG 188

QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 189 ATCTCGGCGACGGGTGGGCACCGCGCGGACGCGGACCTTTCCCGCTTCGCGCTC 248

QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 249 TCGCGCTTCCCGCTGTCGCCCTACGATGTACGAGTACGAGCGCAAGCGCAGAGCGGCCG 308

QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 309 TGCCCGCAGTGCACACTAAGTACAAAGCGCCACAAGGAGGCCACACAGTACACGGTGA 368

QY 81 GluGlyAspAspThrAspAlaAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 369 GAAATAGGATGTGGATGCTGACGATGTGAGTGACTACAACTACCAAGCATCTGGCAAC 428

QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Db 429 CAGGATCAGAAAGATTGCTGAGAGAAATGCTCACCTGGCGGACAAACTCACGTGGC 488
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluileGlyLeuSerLysTyrAsp 140
Db 489 AGT---GATATTGGCTGGCTAAAGTATGACAGCGGTGAAATTGGGCATAGGAAGTATGAC 545
Qy 141 SerGlyGluileProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 546 AGTGGTGAGATCCTCTGTGATATATCCCTGCTACTAACTCATAGCCAGATCTCAGGAGAG 605
Qy 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 606 ATTCTGGAGCTTCCCTGTAT---CATATGATGTCTCTGTGGGAAACATTGGCGGGCGT 662
Qy 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 663 GGACATCAATTTCTTATGTAAATCATTTCTCCAAACCCATCGAGGGAGTTCTCCGGTAGC 722
Qy 200 IleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGly 219
Db 723 CTTGGCAATGTTGCATGGAAGAGAGGGTGGATGGATGGAAATGAAG---GATAAAGT 779
Qy 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
Db 780 GCAATTCCTATGACCAATGGAACAAGCATTTCTCCATCAGAAAGGGCGTGGAGTTGCTGAT 839
Qy 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 840 ATTGATGCTTCTACTGATTATGACATGAAGATGCCTTACTGAATGATGAACCTCGGCAA 899
Qy 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
Db 900 CCTCTATCTAGAAAAGTGCATTTCTTCATCCAGAAATAATCCGTACAGATGGTCATT 959
Qy 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
Db 960 GTGCTAGCTTTGGCTGTCTATGTCATATCTTGGGCTACCGCTATCACATCCCTGTGAAC 1019
Qy 300 AsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluileTrpPheAlaLeuSerTrp 319
Db 1020 AATGCATATCCACTGTGGCTTTTATCGTGCATATGTCGATGAGATCTGGTTGTCTCTCG 1079
Qy 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeu 339
Db 1080 ATTTTGATCAGTTCCTCCAAAGTGTCTCCCAATCAACCGTGAACATACCTTGATAGACTG 1139
Qy 340 AlaLeuArgTyrAspArgGlyGluProSerGlnLeuAlaValAspIlePheVal 359
Db 1140 GCTTTAAGGTATGACCGAGAAGGTGAACCATCTCAATTAGCTCTGTTGATATTTTGTG 1199
Qy 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
Db 1200 AGTACTGTGATCCAAATGAAGGACCTCTCTGTCATCTGCAATACCTGCTTTTCCATC 1259
Qy 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
Db 1260 CTTGCTGTGATATCCGGTTGACAAAGTATCTTGCTATGTTTCGGATGATGGAGCTGCT 1319
Qy 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPhe 419
Db 1320 ATGCTGACTTTGATGCTCTCTGMAACTTCAGAGTTTGTCTAGAAATGGGTTCCGTTTC 1379
Qy 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
Db 1380 TGTAAGAAGTACAAACATAGACCTTAGGGCCCGGAATGGTACTTTGTCTCAGAAAAATTGAT 1439
Qy 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu 459
Db 1440 TACTTTGAAAGCAAAAGTTCAAACCTCATTTGTGAAAGAACCCCGGCGCATGAAGAGAGAA 1499
Qy 460 TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
Db 1500 TATGAGNAATTCAAAGTTCGTATCAATGGTCTTTGTAGCCAGGACAAAAAGTTCCCGAG 1559

Qy 480 GluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 499
Db 1560 GAGGGATGGATCATGCAAGATGTTACACCTTGGCCTGGGAACAATACTAGGGACCATCTCT 1619
Qy 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
Db 1620 GGAATGATTCAAGTTTTCCTGGGTCAAGTGGAGGGCTTCAGCGTTGAAGCAATGAACCTT 1679
Qy 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAla 539
Db 1680 CCTCGTTGGTTTATGTGTCTCGTGAACCAAGCTCTGGATTCCAACATCAAGAAAGGCT 1739
Qy 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
Db 1740 GGTGCCATGAATGCATTTGTTGTTATAGCTGTCTTACTAATGGGCAATACATGTTG 1799
Qy 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
Db 1800 AATCTTGATTGTGACCACTACATCAATAATAGCAAGGCTCTTCGAGAAGCTATGTGCTTC 1859
Qy 580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
Db 1860 CTTATGACCCAAACCTTAGGAAGGATGTCTGTTATGTCCAATTTCTCAGAGGTTTGTAT 1919
Qy 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
Db 1920 GGTATTGATAGGAATGACCATATGCAACAGGAACACTGTGTTTTTCGATATTAACTTG 1979
Qy 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 1980 AGAGTCTTCACGGCATTCAAGGGCCAGTTTATGTGGAACTGGTGTGTGTTTAAACAGA 2039
Qy 640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2040 ACGGCTTATATGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTCTCTCTCG 2099
Qy 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2100 CTTTGTGGGGAAGGAAAAAGCTCAAAATCTAAGAG---AGCTCGGAAAAAGAGAG 2156
Qy 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGly 699
Db 2157 TCATATAGACACGACAGACAGTCTGTACCAAGTATTTAATCTCGAGATATAGAGAAAGG 2216
Qy 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2217 ATTGAAGGTTCTCAGTTTGTATGATGAGAAATCGCTGATTATGTCTCAATATGAGCTTGAG 2276
Qy 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
Db 2277 AAGAGATTTGGCAGTCCAGTGTGTTTGTAGCCCTCTACTCTGATGGAATATGGTGGTGT 2336
Qy 740 ProGlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2337 CCACAACTGCAACTCCAGAGTCTCTTCTGAAGAAGCTATTATCATCATCAGCTGTGGC 2396
Qy 760 TyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu 779
Db 2397 TATGAGGACAAAACGACTGGGAACTGAGATTGGGTGGATATATGTTCTCTGTTTACAGAA 2456
Qy 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro 799
Db 2457 GACATTTCCCGGATTCAGAGTATGCTCGAGGCTGGCGATCAATCTACTGTGATGCCT 2516
Qy 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2517 AAGGACACGCTTTCAGAGGATCTGCTCTATCAACCTTTCCGATCTGTTGAATCAAGTG 2576
Qy 820 LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr 839
Db 2577 CTTTCGGTGGCTCTTGTTCATTTGAATTTCTTTTCAGGAGGCAATTTGCCCATATGTAT 2636


```
QY 840 GlyTyrGlyGlyArgLeuLeuPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2637 GCCTATGAGGCGCGCTAAATTCCTGGAGAGATTTCCTTATCAACACAACTTAT 2696
QY 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2697 CCACCTCACATCAATCCGCTCCTCTGTACTGTCATATTGCCAGCAGTTTGTCTTTCACT 2756
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPhe 899
Db 2757 GGAAGATTTCATCATCCCAAGATTAGTAACCTAGAGAGTGTGTGGTTTATATCGCTCTTT 2816
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
Db 2817 ATCTCAATCTTTGCCACTCGTATCCTTTGAGATGAGGTGGAGTGTGTGGCATGATGAA 2876
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 2877 TGGTGGAGGAACGAGCAGTTCCTGGGTCAATTGGTGTGATTTCTGCCGCAATTATTTGCCGTC 2936
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 2937 TTCCAGGGTCTCTGAGGTGTTGCTGGTATCGACAGAGCTTCACTGTCACTCTAAG 2996
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 2997 GCCACTGACGAAGAAGTGATTTTGGCGAGCTTACATGTTCAAGTGGACAAACGCTTCTG 3056
QY 980 IleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3057 ATCCACCAACCACTATTTTGTATCATCAACCTGGTGGCGTGGTCTGGCAATTTCTTAC 3116
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3117 GCATTCATAGCGGTACAGTCATGGGGACCTCTTTCCGGAGACTCTTCTTTCGGTTC 3176
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3177 TGGGTGATTTGCACCTGTACCCCTTCTCAAGGGCCTCATGGGAAGCAGAACCGCAGC 3236
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3237 CGACCATTTGTCTGTCTGGGCTATCCTCTTGGTGTGATCTTTTCCCTGTATGGGGTT 3296
QY 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3297 CGTATCGATCCATTACCAACCCGGGTCACTGGCCCTGATATCCGAAATGTGGCATCAAC 3356
QY 1080 Cys 1080
Db 3357 TGC 3359
```

```
RESULT 15
US-10-437-963-39762/c
; Sequence 39762, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 39762
; LENGTH: 4264
```

```
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MPT4530_4326C.1
US-10-437-963-39762

Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 5107.50 Matches: 944
Percent Similarity: 93.94% Conservatives: 64
Best Local Similarity: 87.98% Mismatches: 51
Query Match: 88.40% Indels: 15
DB: 17 Gaps: 4

US-09-900-237A-30 (1-1080) x US-10-437-963-39762 (1-4264)

QY 9 LysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGlyLeuGlyThr 28
Db 3431 AAGTCGGGAGGACGCGGAGCGCGGCGTCCAGATCTTAGGTGACGAG--GGC--- 3377
QY 29 ThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProValCysArgPro 48
Db 3376 -----GACGTCTGGGGTTCCCGGTGGCCGGCCA 3347
QY 49 CysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLysThrLysTyr 68
Db 3346 TGCTACGAGTAGCAGCGCAGGAGCGCGCAGCGCGTSCCCGCGAGTGCACCAAGTAC 3287
QY 59 LysArgHisArgGlySerProAlaIleArgGlyGluGluGlyAspAspThrAspAlaAsp 88
Db 3286 AAGCCCAACAAAGGGAGCGCCCGGATCTTGGGGATGAAAGCGATGATGTTTCATCGGAT 3227
QY 89 AspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspGlnLysGlnLysIleAla 108
Db 3226 GATGCTAGTGTGATGTAATCTCCAACTCCGCAACAGGACCAATAGCACCAAGATTGCC 3167
QY 109 AspArgMetArgSerTrpArgMetAsnThrGlyGlySerGlyAsnValGlyHisProLys 128
Db 3166 GAGAGGATGCTCACCTGGCGCATGAATCTGGGAGGAATGATGATATTGCTCATTAAG 3107
QY 129 TyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIleProArgGlyTyr 148
Db 3106 TATGATAGTGTGTGAGATTGGTTCATCCCAAGTATGACAGTGGTGAATCCCTCGCATATAT 3047
QY 149 ValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyValAserProAspHis 168
Db 3046 ATCCCATCGTCACTCACAGCCAGATCTCAGGTGAATTCCTGGAGCGTCCCTCAT--- 2990
QY 169 HisMetSerProThrGlyAsnIleSerArgArgAla---ProPheProTyrValAsn 187
Db 2989 CATATGATGTCTCCCGTTGGGAACATTGGCAGACGTTGGGCATCCATTTCCCTATGTGAAC 2930
QY 188 HisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValAlaTrpLysGlu 207
Db 2929 CATTCACCAACCCCATCAAGGAGTTCCTGTAGCTTGGCAATGTTGTCATGGAAAGAG 2870
QY 208 ArgValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThr 227
Db 2869 AGAGTAGATGCTGGAAATGAAG---GATAAAGTGCATTTCCCATGGCTAATGGTACT 2813
QY 228 SerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsn 247
Db 2812 AGCATTTGCTCTTCAGAAAGTTCGTGGGTTCGTGACATTGATGTCATCTACTGACTATAAC 2753
QY 248 MetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIle 267
Db 2752 ATGGAGGACGCTTACTGAATGATGAACACGCCCAACCTCTATCGAGAAAGTGCACAAAT 2693
QY 268 AlaSerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuValLeuSer 287
Db 2692 TCGTCATCCAGATAAATCCGTACAGGATGTTATTGTGCTCCGATTCATGTTCTTATGT 2633
QY 288 IlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeu 307
```


Db 2632 ATATCTTCACATACCGTATACAAATCTCTGCGCTAATGTCATACCCGCTATGTTGCTC 2573
QY SerValIleCysGluLeuTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrp 327
Db 2572 TCTGTCAATGTGAGATTGTTGTTGCTCTATCTGCTGATTTCTGGATCAGTTCGCCAAGTGG 2513
QY PheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGly 347
Db 2512 TCCCCAATCAACCGTGAACACTTACCTGGATAGGCTGGCTTTAAGGTATGACCGGGAAGGT 2453
QY GluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGlu 367
Db 2452 GAACCAATCAATGGCTCTCTGTGACATTTTGTCTGAGTCTGTGGATCCCTATGAAGAAGA 2393
QY ProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAsp 387
Db 2392 CCTCCTCTTGTCACTGCCAATACTGTGCTCTCCATCCTTGTCTGGATTACCCCTGTTGAT 2333
QY LysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAla 407
Db 2332 AAGGTATCTTGTCTATGTCTGTATGATGGAGCTGCAATGTTAACTTTGATGCGCTGTCT 2273
QY GluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluPro 427
Db 2272 GAGACTTCAGATTGCTAGAAAATGGGTACCGTTCGTAAAAAGTACACATAGAACCC 2213
QY ArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnPro 447
Db 2212 AGAGCTCCAGAGTGGTACTTTTGCACAGAAAATTTGATTACTTGAAGACAAAGTTCAGGCT 2153
QY SerPheValLysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIle 467
Db 2152 TCTTTTGTAAAGATCGTCTGTCATGACAGAGGAATATGAAGAATTTAAAGTTTCGTGTT 2093
QY AsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGly 487
Db 2092 AATGCTCTGTCGCCAAGGACAGAAAGTTCCTGAGAAAGATGGATTATGCAAGATGCC 2033
QY ThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGly 507
Db 2032 ACACCTTGGCTCGTAAACAATACAGAGACCATCCTGGAATGATTACAGTTTTCCTTGGT 1973
QY HisSerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArg 527
Db 1972 CATAGTGGAGGCTTGATACCGAAGCAATGAGCTTCCTGTTTAGTCTATGTGCTCGT 1913
QY GluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArg 547
Db 1912 GAGAAACGTCTGGATTCCAAACACCATAAAAAGGCTGGTCTATGAATGCATGTTGCTGT 1853
QY ValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIle 567
Db 1852 GTATCAGCTGTCCTTACTAATGGACAATACTTGTGAATCTTGACTGTGATCACTACATC 1793
QY AsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyPro 587
Db 1792 AACAAATAGCAAGCTCTCCGAGAGGCTATGTGCTTCTTATGGATCCAAACCTAGGAAGG 1733
QY GlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyr 607
Db 1732 CGTGTCTGTATGTCCAATTCCTCAGAGGTTTGACGGTATCGATGAGNATGATCGATAT 1673
QY AlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGly 627
Db 1672 GCACACAGGAACACCGTGTGTTTTCGATATTAATTTGAGAGTCTCGATGTTCTCCAAAGA 1613
QY ProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluPro 647
Db 1612 CCAGTTTATGTGGAAACCGTGTGTGTTCAACAGAACAGAGCTCTTATGGTTATGAACCC 1553
QY ProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAla 667

Db 1552 CCCATTAGCAGACAGAGGCCAGGTACTTCTCTTCGCTTTGTGGGGACGAAAGAAGACA 1493
QY SerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSer 687
Db 1492 AAAAAGTCCAAAGAGAGACGACGGAAGAAAAGTACACAAACATGTGGACAGTTCT 1433
QY ValProValPheAsnLeuGluAspIleGluGlyValGluGlyAlaGlyPheAspAsp 707
Db 1432 GTGCCAGTGTTTAACTTCGAAGATATAGAGGAAGGATTTGAGTTCTGATTTGATGAT 1373
QY GluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAla 727
Db 1372 GAGAAATCACTACTGATGTCTCAATCAGCTTAGAGAAAAGATTTGCTCAATCTAGTTGT 1313
QY PheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSer 747
Db 1312 TTTGTAGCTCCACTCTGATGGAATATGGTGGTGTCTCTCAATCCCAACTCCAGAACT 1253
QY LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTrpGly 767
Db 1252 CTCTGAAGAGGCCATACATGTTATCAGCTGTGGCTATGAACACAAAAGTACTGGGA 1193
QY ThrGluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMet 787
Db 1192 ACTGAGATTGGATGGATCTATGGTCTCTGTACAGAAGATATCTCACTGGATTCAAGATG 1133
QY HisAlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySer 807
Db 1132 CATGCAAGTGGCTGGCGGTCAATCTACTGATGCCCTAAGCCAGCAGCATTTCAAGGGTCT 1073
QY AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 827
Db 1072 GCCCAATTAATCTTTCCGATCGTCTGAACAGGTGCTTCGGTGGGCTCTTGGTTCTGTG 1013
QY GluIleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPhe 847
Db 1012 GAAATCTTTTACGAGGCATTTGCCATATGTTATGGGTACGAGAGACGGCTTAAGTTC 953
QY LeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeu 867
Db 952 CTGAGAGATTGCTTTACATCAACACCATTTTATCCACTCACATCGATTCACCTCTC 893
QY ValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIle 887
Db 892 TTGTACTGCATATTTGCCAGCTATCTGCTCTCTCCTGGGAAGTTTATCATACACAGATT 833
QY SerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIle 907
Db 832 AGCAACTTTGCGAGTATTTGGTTTATCTCTCTCTTTCTGTCAATCTTTGTACTGGTATC 773
QY LeuGluMetArgTrpSerGlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrp 927
Db 772 CTTGAGATGAGGTGAGTGTGTGGTATCGATGAATGGTGGAGGAATGACAGTTTGG 713
QY ValIleGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeu 947
Db 712 GTTATTGGAGTATTTCCGCTCATCTATTTCGGTCTTCCAGGCTCTGTCTGTAAGTCTCT 653
QY AlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPhe 967
Db 652 GCTGGAATTGATACAGCTTCACTGTCCACTCAAAAGGCCTCTGATGAAGAAGCGGATTTC 593
QY AlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIle 987
Db 592 GCTGAGCTCTACATGTTTCAAGTGAGCAACCTTCTGATCCCGACCGACCATTTTTCATC 533
QY IleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSer 1007
Db 532 ATCAACTGTGTCGGTGTGTGTTGCTGATCTCTCCACGCTATCAACAGTGGCTACAGTCA 473
QY TrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeuTyrPro 1027
Db 472 TGGGCTCCGCTCTTTGGGAAGCTCTTCTTTCCTTCTGGGTGATGTTGCCACCTGTACCCC 413

```
QY      1028 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTnpAla 1047
Db      412 TTCCTTAAGGGTCTCATGGTCGGCAGAACCGTACACCCGACCATGTTGTTGTTGGGCC 353

QY      1048 ValLeuLeuAlaSerIlePheSerLeuLeuTnpValArgValAspProPheThrThrArg 1067
Db      352 ATCCTCCTTGCATCGATCTTCTCCTTGCTGTGGTTGCTGATCCATTCACCCCGT 293

QY      1068 LeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db      292 GTCACAGGCCCTGATACCCCAAAATGTGTATCAACTGC 254
```

Search completed: August 23, 2004, 01:08:01
Job time : 1258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 00:36:19 ; Search time 29 Seconds
(without alignments)
3582.305 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVCO.....VDPFTTLAGPNITCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	4507	78.0	1065	T52054	cellulose synthase
2	3963.5	68.6	1081	T05351	cellulose synthase
3	3801	65.8	1026	T51579	cellulose synthase
4	3778.5	65.4	1065	F84649	probable cellulose
5	3666.5	63.5	1084	T08583	cellulose synthase
6	3583	62.0	1081	T52028	cellulose synthase
7	3579	61.9	1088	H84604	probable cellulose
8	3477	60.2	974	T10797	cellulose synthase
9	3302.5	57.2	958	T04870	cellulose synthase
10	2816	48.7	685	T10800	cellulose synthase
11	2315	40.1	1181	D86157	hypothetical prote
12	2257	39.1	1111	T05646	hypothetical prote
13	2220.5	38.4	1145	T51546	cellulose synthase
14	2212.5	38.3	1036	D84741	probable cellulose
15	2163.5	37.4	979	C86446	probable cellulose
16	1941.5	33.6	583	T02209	cellulose synthase
17	1094	18.9	757	T02561	probable cellulose
18	1076	18.6	712	T02552	cellulose synthase
19	1057.5	18.3	755	T02553	cellulose synthase
20	1034.5	17.9	748	T02560	cellulose synthase
21	983	17.0	689	T08918	hypothetical prote
22	963.5	16.7	686	T08919	hypothetical prote
23	959.5	16.6	727	T08920	hypothetical prote
24	943.5	16.3	828	E71417	hypothetical prote
25	872.5	15.1	710	B71417	hypothetical prote
26	430	7.4	693	A82275	cellulose synthase
27	289.5	5.0	759	D70422	cellulose synthase
28	271.5	4.7	322	T12093	TGACG-motif bindin
29	270	4.7	326	T08591	TGACG-motif bindin

30	264	4.6	874	2	AB0985	probable polysacch
31	259.5	4.5	1596	2	T31338	cellulose synthase
32	258	4.5	888	2	E91180	probable cellulose
33	257	4.4	692	2	S47754	hypothetical prote
34	257	4.4	888	2	H65151	hypothetical 101.6
35	255.5	4.4	729	2	A98320	cellulose synthase
36	255.5	4.4	729	2	AD2963	cellulose synthase
37	255.5	4.4	861	2	T39714	cellulose synthase
38	251.5	4.4	888	2	G86026	probable cellulose
39	239.5	4.1	322	2	T08592	TGACG-motif-bindin
40	238.5	4.1	754	2	A43735	bcsA protein - Ace
41	232	4.0	664	2	A95889	probable cellulose
42	207	3.6	723	2	C36963	cellulose synthase
43	179	3.1	768	2	B97083	glycosyltransferas
44	164	2.8	654	2	AB2932	beta 1,3 glucan sy
45	164	2.8	654	2	D98350	hypothetical prote

ALIGNMENTS

RESULT 1

T52054

cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000

C:Accession: T52054

R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H

A:Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.

A:Reference number: Z13745; MUID:98111412; PMID:9445479

A:Accession: T52054

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1065 <ARI>

A:Cross-references: EMBL:AF027174; PIDN:AAC39336.1

C:Genetics:

A:Note: Ath-B

C:Function:

A:Description: EC 2.4.1.-; cellulose synthase [validated, MUID:98111412]; involved in as

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 78.0%; Score 4507; DB 2; Length 1065;

Best Local Similarity 78.4%; Pred. No. 0;

Matches 836; Conservative 98; Mismatches 107; Indels 26; Gaps 8;

Qy 19 CQICADGLGTTLDGVDFTACDVCRPCPCYERKEGTOACLCQCKTKYKRRHSGSPAIR 78

Db 20 CQICSDNVGKTVGDGRFVACDICSFPVCRPCYERKDGNSQCPQCKTRYKRLKSGSPAIP 79

Qy 79 GEEGDDTDADGS-DFNYPASGTDQKKTADNRSHRMWITGGSGNVGHPKYSDEIGLS 137

Db 80 GKDDEGLADEGTVFEFNYP-----QKEKTSERMLGWHLTRGKGEMGEPOYDK----- 127

Qy 138 KYDSGEIPIRGVPSVTNSQ-MSGBI PGASPDHMHMSPGNSIRRAPPY-VNHSNPSRS 195

Db 128 -----EVSHNHLPLRTSQDTSGEFSAASPERLSVSSTIAGKKRLPYSSDVNQSPNRIV 182

Qy 196 FSGSIGNVAMKERYDVGWKMQDKGAIPTMTGTSTIAPSEGRAATIDASTEYNMEDALLND 255

Db 183 DPGVLGNVAMKERYDVGWKMQDKNTGV---SITQAASE-RGGVDIDASTDILADEALLND 238

Qy 256 ETROPLSRKVPASSKINPRMIVLRLVLSLFLHYRLNPNVNPVNAFLMLLSVICIWF 315

Db 239 EARQPLSRKVPASSSRINPRMIVLRLVLCFLHYRITNPVNPVNAFLMLVSVICIWF 298

Qy 316 ALSMILQDFPKWPPINRETYLDRLALRYDREGESQLAAVDIFVSTVDLKEPPIVTANT 375

Db 299 ALSMILQDFPKWPPVNNRETYLDRLALRYDREGESQLAAVDIFVSTVDLKEPPIVTANT 358

Qy 376 VLSILAVDYPVDKVCYVSDGDGSMLTFDALAESEFARKWVPFKKYDIEPRAPEFYFC 435

Db 359 VLSILAVDYPVDKVCYVFDGGAAMLFSFESLAETSEFARKWVPFKKYDIEPRAPEYFA 418

QY	436	OKIDYLDKQVQPSFVKDRAMREYEEFKIRINALVSKALKVPEEGWIMQDGTTPGNNNT	495
Db	419	AKIDYLDKQVQPSFVKDRAMREYEEFKIRINALVSKALKVPEEGWIMQDGTTPGNNNT	478
QY	496	RDPGMIQVFLGSHGGLDTEGNEPLRLVTVSREKPGFOHHKKAGAMNALVRVSAVLITNG	555
Db	479	GDHPGMIQVFLGSHGGLDTEGNEPLRLVTVSREKPGFOHHKKAGAMNALVRVSAVLITNG	538
QY	556	QYMLMLDCDHYINNSKAVREAMCFLMDPNLGRQVQVCVQPQFPDGLDRNDRYANRNTVFF	615
Db	539	PFILMLDCDHYINNSKALREAMCFLMDPNLGRQVQVCVQPQFPDGLDKNDRYANRNTVFF	598
QY	616	DINLRLGLDGIQGPVVVGTGCVFNRTAIGVEPIKA--KKPGFLASLCGGKKKASKSKKR	673
Db	599	DINLRLGLDGIQGPVVVGTGCVFNRTAIGVEPIKVKVHKPKSLLSKLGGSRKKSARKK	658
QY	674	SSDKKSKNKHVDSPVFVFNLEDEEGVEGAGDDEKSVLMSQMSLEKRFQSAFVASTL	733
Db	659	ESDKKSKGRHTDSTVPFNLDDIEEGVEGAGDDEKALLMSQMSLEKRFQSAFVASTL	718
QY	734	MEYGVQPSSTPESLLKEAHHVISCYEDKSEWGTETIGWISVTEIDLTGFKMHARGWR	793
Db	719	MENGGVPPSATPENLLKEAHHVISCYEDKSDWGEIGWISVTEIDLTGFKMHARGWR	778
QY	794	SVYCKPKRPAFKGAPINLSDRNLQVLRWALGSVELFSRHCPLMVYGGRLKELERPAY	853
Db	779	SYICMPKLPFAFKGAPINLSDRNLQVLRWALGSVELFSRHCPITWYNGRLKFLERPAY	838
QY	854	INTTIYPLTSLPLLYCYILPAICLTGKFMPEISNLASIWFIALFLSIFATGILEMRWS	913
Db	839	VNTTIYPTITSIELLMYCTLLAVCLFTNQPIIQISNIASIWFLSLFSIFATGILEMRWS	898
QY	914	GVGIDEMWRNEQFWIGGISAHLPFAVQGLLKVLAGIDTNFTVTSKANDEBGPFAELIYMF	973
Db	899	GVGIDEMWRNEQFWIGGVSAHLFAVQGLILKVLAGIDTNFTVTSKASDEGDFAEIYLF	958
QY	974	KWTTLLIPPTTILLINMGVWAGTSVAINSGVQSGWGLFGKLFPAFWIIVHLIYPLFKGLM	1033
Db	959	KWTTLLIPPTTILLINMGVWAGTSVAINSGVQSGWGLFGKLFPAFWIIVHLIYPLFKGLM	1018
QY	1034	GRQNRTPTIIVWAVLLASIFSLMLWVRVDPFTTRLAGPNIQTCGINC 1080	
Db	1019	GRQNRTPTIIVWVLLASIFSLMLWVRIDPFTSRVTGPDILECGINC 1065	
RESULT 2			
T05351			
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana			
N:Alternate names: protein F8B4.110			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999			
C:Accession: T05351			
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De			
ewes, H.W.; Mayer, K.F.X.; Schueller, C.			
submitted to the Protein Sequence Database, February 1999			
A:Reference number: Z15409			
A:Accession: T05351			
A:Molecule type: DNA			
A:Residues: 1-1081 <REV>			
A:Cross-references: EMBL:AL034567			
A:Experimental source: cultivar Columbia; BAC clone F8B4			
C:Genetics:			
A:Gene: RSW1			
A:Map position: 4			
A:Intron: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1; 76			
C:Keywords: glycosyltransferase; hexosyltransferase; P-loop			
Query Match 68.6%; Score 3963.5; DB 2; Length 1081;			
Best Local Similarity 68.0%; Pred. No. 3.2e-282;			
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;			

QY	2	DGDADALKSGRHAGDVCOICADGLGTTLDGVTACDVCVRFPVPCPYEHERKEGTQAC	61
Db	25	DGDTPLKNNM--QOICQICGDDVGLAETGDFVACNECAPVCPVPCPYEYERKDTQCC	81
QY	62	LOCKTKYKRRHSPAIRBEGDDTDADGSDGFNYPASGTEDQOKTIADRMRSWRMTGGS	121
Db	82	POCKTRFRHRGSPRVEGDEDDDDIENEENYAQGANKARHQRHGEFSS-----	133
QY	122	GNVGHPKYDSGIGLSKYD--SGEIPRGYVPSVINSQMSGELPGASPDHMMKSPGNTLS	178
Db	134	----SSRHESQIPILLTHGHTVSGEIRTPDTSVARTT--SGPL-GPSDRNALSSP--YID	184
QY	179	RRAPPYVNNHSPNPREF--SGSIGNVAWKERVDGWMKQDKGAIPTWNGTSTIAPSGRAA	237
Db	185	PRQPVV--RIVDPKOLNSYGLGNVWKVERVEGWKLKQEKMLQMTG---KYHEGKGG	238
QY	238	TDIDASTEYNMEDALLNDETROPLSRKVPDIASKINPYRMVILVRLVLVLSIFLHVLTNP	297
Db	239	-EIE-GTGSNGBELQADDTLRLPMSRVVPIPSRLTFYRVVILRLIILCLFPLQVTRTHP	296
QY	298	VNAYPLMLLSVCEIWEFALSWILQDPKWFPIINRETYLDRLALRYDREGSPSOLAAVDI	357
Db	297	VKNAYPLMLTSVCEIWEFAPSWLLQDPKWPINRETYLDRLAIRYDRDGSPSOLVFPDV	356
QY	358	FVSTVDPLEKPEPIVTANTVLSILAVDYPVDKVCVSDDGASMLTFDALAETSEFARKW	417
Db	357	FVSTVDPLEKPELVANTVLSILSDVDYPVDKVCYVSDDGASMLTFESLSEAEFAKKW	416
QY	418	PFVKYVDIEPRAPBFYQKIDYLDKQVQPSFVKDRAMREYEEFKIRINALVSKALKV	477
Db	417	PFCKNFNIEPRAPBFYQKIDYLDKQIQPSFVKERRAMREYEEFKIRINALVAKAQKI	476
QY	478	PEEGWIMQDGTTPGNNTRDHPGMIQVFLGSHGGLDTEGNEPLRLVTVSREKPGQHHK	537
Db	477	PEEGWIMQDGTTPGNNTRDHPGMIQVFLGSHGGLDTEGNEPLRLIIVSREKPGQHHK	536
QY	538	KAGAMNALVRVSAVLITNGQYMLNLCDDHYINNSKAVREAMCFLMDPNLGPQVCYQFPQR	597
Db	537	KAGAMNALIRVSAVLITNGAYLLAVDCDHYFNNSKAIKEAMCFMMDPAIGKKCCYQFPQR	596
QY	598	FGIDNRNRYANRNTVFFDINRLGDLGQGPVYVGTGCVFNRTAIGVEPPPIKAK--KFG	655
Db	597	FGIDILHRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFNRTAIGVEPPPIKAK--KFG	656
QY	656	FLA-SICGKKKASKKSSDKKSNKHVDSPVFVFNLEDEEGVEGAGDDEKSVLMS	714
Db	657	IIVKSCGSRKKGKSKSKTYNEKRRKNSDNAPLFNMEDIDEGF--GYDDERSILMS	714
QY	715	QMSLEKRFQGSAAFVASTIMEYGVQPSSTPESLLKEAHHVISCYEDKSEWGTETIGWY	774
Db	715	QRSVEKRFQGPSVFIAATFMEQGIPTTNPATLLKEAHHVISCYEDKTEWGEIGWIY	774
QY	775	GSVTEDILTGFQKHARGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILLFSRH	834
Db	775	GSVTEDILTGFQKHARGWISYICNPRPAPKGSAPINLSDRNLQVLRWALGSVEILLFSRH	834
QY	835	CPLWYGYGRLKFLERFAVINTTIYPLTSLPLLYCYILPAICLTGKFMPEISNLASTW	894
Db	835	CPWIYGYGRLKFLERIAVINTVIPTSIPUIACILPAICLTGKFMPEISNLASTW	894
QY	895	FIALFLSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLPFAVQGLLKVLAGIDTNF	954
Db	895	FILLFISIAVTGILELRWSGVGIDEMWRNEQFWIGGTSAHLFAVQGLLKVLAGIDTNF	954
QY	955	TWTSKANDEBGPFAELYMFKWTLLIPPTTILLINMGVWAGTSVAINSGVQSGWGLFGK	1014
Db	955	TWTSKANDEBGPFAELYIFKWTALLIPPTTILLINMGVWAGTSVAINSGVQSGWGLFGK	1014
QY	1015	LEFPAFWIIVHLIYPLFKGLMGRQNRPTTIIVWAVLLASIFSLMLWVRVDPFTTRLAGPN	1072
Db	1015	LEFPAFWIIVHLIYPLFKGLLGRQNRPTTIIVWVSVLLASIFSLMLWVRINFPVD--ANPN	1070

QY	348	EPSQLAAVDIFVSTVDPLKEPPIVTANTVLSILAVDYPVDVKVSCYVDDGASMLTFDALA	407
Db	334	EPSQLAPVDVFTVDPMPKEPLVTANTVLSILAVDYPVDVKVSCYVDDGASMLTFEALS	393
QY	408	ETSEFARKWVPVKKYVDIEPRAPFVFCQKIDYLDKQVQPSFVKORRAMKRYEYFKIRI	467
Db	394	ETAEFSSKWVPFCKFNIEPRAPFVFSQKIDYLDKQVQPSFVKORRAMKRYEYFKVRI	453
QY	468	NALVSKALKVPEEGIMQDGTWPGNNTDHPGMIQVFLHSGGLDTEGNEPLRLVYYSR	527
Db	454	NILVAKAOKIPEDGTMEDGTSWPGNPNRDPHFGMIQVFLHSGGLDTEGNEPLRLVYYSR	513
QY	528	EXRPGFQHHKAGAMNALVRVSAVLTMGOYMLNLDCHYINNSKAVREAMCFIMPDLNLP	587
Db	514	EXRPGFQHHKAGAMNALIRVSAVLTMGAYLLNVDCDHYFNNSKALKEAMCFMPDPAIGK	573
QY	588	QVCYVQFPQRPQIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNFTALYGEYP	647
Db	574	KCCYVQFPQRPQIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCCFNRQALYGYDP	633
QY	648	PIKAK--KPGFTASLCGGKKKASKK--RSSDKKSNKHVDSSVPVFNLEDIEEGVEGAG	704
Db	634	VUTEEDLENIIVKSCFGGRKKGKSKKIPNVEDNRSIKRSQNSVPLFNMEDIDEDVE--G	691
QY	705	FDEKSVLMSQMSLEKRFQCSAAFAVASTLMVEYGGVPPQSSTPESLLKEAIIHVISCGYEDKS	764
Db	692	YEDENSLVASQKRLKRFQCSVPVIAATFMEQGLPSTTNPLTLKKEALHVISCGYEAKT	751
QY	765	EWGTEIGWYGVSTEDILTGFQKHARGWRSVYCMKPRPAFKGAPINLSRLNQLRWAL	824
Db	752	DMGKEIGWYGVSTEDILTGFQKHARGWRSIYCVSPRPAFKGAPINLSRLNQLRWAL	811
QY	825	GSVEILFSHCHPLWGYGGLAKFLERFAVINTIYPLTSLPLLVLCILPAICLLTGKFTM	884
Db	812	GSIEILLRHCPIWGYNGRLKLLERKIAVINTIYPTISPLLAYCMLPAFCLLTNTPFI	871
QY	885	PEISNLASIFIALFSLPATGILEMRWSGVGIDENWRNEQFWVIGGISAHLFAVFOGLL	944
Db	872	PEISNLASLCFMLLFASIVASALELKWSDALEDWRNEQFWVIGGISAHLFAVFOGLL	931
QY	945	KVLGIDNTFTVTSKANDEGFAELYMFKWTLLIPPTTILIIINWGVVAGTSTAINSG	1004
Db	932	KVFAGIDNTFTVTSKASDGDGFAELYVFKWTSLLIPPTTILIIINWGVVAGVSAINSG	991
QY	1005	YQSWGFLFKLPFAFWIVHLYPFLKGLMGQRNRPPTIIVVAVALLASTFSLWVRVDFP	1064
Db	992	YQSWGFLMGKLLFAFWVVAHLYPFLKGLMGQRNRPPTIIVVAVALLASTFSLWVRINFP	1051
QY	1065	FTRLAGPNITQCGI 1078	
Db	1052	-----VSTGTG 1057	
RESULT 5			
T08583			
cellulose synthase (EC 2.4.1.-) catalytic chain - Arabidopsis thaliana			
N:Alternate names: protein T22F8.250			
C:Species: Arabidopsis thaliana (mouse-ear cross)			
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999			
C:Accession: T08583; T09014			
R:Bevan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May			
submitted to the Protein Sequence Database, May 1999			
A:Reference number: Z16442			
A:Accession: T08583			
A:Molecule type: DNA			
A:Residues: 1-1084 <BEV>			
A:Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.250			
A:Experimental source: cultivar Columbia; BAC clone T22F8			
R:Aricoli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H			
Science 279, 717-720, 1998			
A:Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.			
A:Reference number: Z13745; M0ID:98111412; PMID:9445479			

A:Accession: T09014			
A>Status: translated from GB/EMBL/DBDJ			
A:Molecule type: mRNA			
A:Residues: 1-1084 <ARI>			
A:Cross-references: EMBL:AF027173; NID:G2827140; PIDN:AAC39335.1; PID:G2827141			
A:Experimental source: cultivar Columbia			
C:Genetics:			
A:Gene: ATSP:T22F8.250; ACh-A			
A:Map position: 4			
A:Introns: 27/3; 156/3; 191/1; 254/2; 343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3; 819/3; 823/3			
C:Keywords: cell wall synthesis; glycosyltransferase; hexosyltransferase			
Query Match 63.5%; Score 3666.5; DB 2; Length 1084;			
Best Local Similarity 62.4%; Pred. No. 2e-260;			
Matches 687; Conservative 164; Mismatches 191; Indels 59; Gaps 16;			
QY	1	MDGDADA-LKSGRHGAGDVCCIADGLGTTLDGVDFTACDVCRCFPVCYHERKEGTQ	59
Db	20	INADESARIRSQVELSGQTCQICGDEIELTVSELFAVACNECAFPVCRFCYERREGNQ	79
QY	60	ACLOCKTKYKRHRGSPAIRGEEGDDTDADGSDDFNYPASGTEDQKQIADRMRSWRMNTG	119
Db	80	ACPQCCKTRYKRIKGSFV---DGDDEBEDIDDLVEYFDHGMV-PEHAAEALSSRLNTG	135
QY	120	GSQNVGHPKYSGBEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMMSP--TGN	177
Db	136	RGGLDSAP-----PGSQIPLLTVCDEADW--YSDRHALLIVPSTGYG	176
QY	178	SRRAFPFVYVNSPNFS-----REFSGSIGNVANKERVDGKMKQ-DKGALPMTNG	226
Db	177	NRVYFAPFTDSSAPPQARSMPQKDIABE--GYGSVANKORMEVWKRKRGKQLVTKHEG	234
QY	227	TSIAPSEGRGAATDIDASTEYNNMEDALLNDETQPLSRKVPKVIASSKINPYRMVILRVVL	286
Db	235	-----GNGREGSNDDB---ELDDPDMPMWDEGQPLSRKLPTRSSRINPYRMILLCETAIL	287
QY	287	SIFLHYRLTNPNVRNAYPLWLLSVICEIFALSWILDDQFPKVPFPIINRETYLDRLALYDRB	346
Db	288	GLFFHYRILHPVNDAYGLWLTSTVICEIFAVSWILDDQFPKVPPIERETVLDRLSLRYEKE	347
QY	347	GPESOLAADVIFVSTVDPLKEPPIVTANTVLSILAVDYPVDVKVSCYVSDGASMLTFDAL	406
Db	348	GKPSGLAPVDVSVTVDELPEKPLITANTVLSILAVDYPVDVKVACVSDGGAAMTFEAL	407
QY	407	ASTSEFARKWVPVKKYVDIEPRAPFVFCQKIDYLDKQVQPSFVKORRAMKRYEYFKIR	466
Db	408	SITAEFARKWVPFCKFNIEPRAPFVFSQKIDYLDKQVQPSFVKORRAMKRYEYFKVYK	467
QY	467	INALVSKALKVPEEGIMQDGTWPGNNTDHPGMIQVFLHSGGLDTEGNEPLRLVYYS	526
Db	468	INALVATAQKVPEEGIMQDGTWPGNNVDRHPGMIQVFLHSGVRDTEGNEPLRLVYYS	527
QY	527	REKRPGFQHHKAGAMNALVRVSAVLTMGOYMLNLDCHYINNSKAVREAMCFMPDNLG	586
Db	528	REKRPGFQHHKAGAMNSLIRVSAVLSNAPYLNVDCDHYINNSKAIRESMCFMMDPQSG	587
QY	587	POVCYVQFPQRPQIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGE	646
Db	588	KVCYVQFPQRPQIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFRRQALYGF	647
QY	647	PIPKAKPGFIAS-----LCGGKKKASKSKRSDDKKSNKHVDSSVPVFNLEDIEBG	699
Db	648	APKKKKPGKTCNCWPKWCCCLCGLRK--KSKTKAKDKTKNTK--ETSKQIHALENVDRG	703
QY	700	VBGAGFDDDEKSVLMSQMSLEKRFQCSAAFAVASTLMVEYGGVPPQSSTPESLLKEAIIHVISCG	759
Db	704	VIVFVSNVEKSEATQLKLEKFGQSPVFAVSAVLQNGGVPRNASPACLLRAIQUIVSCG	763
QY	760	YEDKSEWGTGTEIGWYGVSTEDILTGFQKHARGWRSVYCMKPRPAFKGAPINLSRLNQL	819
Db	764	YEDKTEWKGTEIGWYGVSTEDILTGFQKHARGWRSVYCMKPRPAFKGAPINLSRLNQLH	823
QY	820	LRWALGSEVILFSRCHPLWGYGGLAKFLERFAVINTIYPLTSLPLLVLCILPAICLLT	879

```

Db      824 LRWALGSVEIFLSRHCHPIWYGGGLKWLRFYSVNSVVPWISLPLIVCSLPAVCLLT 883
Qy      880 GKTIPIPSINLASIWFIALFSLFATGILEMRSGVGIDSEWRNEQFWITGGISAHLPFAV 939
Db      884 GKFTVPEISNYAGILFIMFISIAVTGILEMVGWGVDWWRNEQFWITGGASHLFAL 943
Qy      940 FOGLLKVLAGIDNFTVTSKANDEEGFAELYFKFWITLIPPTILLINMGVVACTSY 999
Db      944 FOGLLKVLAGVNTNFTVTSKAAD-DGAFSELYIFKWTLLIPPTILLINMGVIGVSD 1002
Qy      1000 AINSYGOSWGLFGLKFAFVWVHLVFLKGLMGRQNRTPPTIVWAVLLASIFSLWV 1059
Db      1003 AINSYGDSWGLFGLKFAFVWVHLVFLKGLMGRQNRTPPTIVWAVLLASIFSLWV 1062
Qy      1060 RVDPTFTRLAGPNIOQTGNC 1080
Db      1063 RVNPFVAK-GGPVLEICGLNC 1082

RESULT 6
T52028
cellulose synthase [imported] - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52028
R;Joshi, C.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z25890
A;Accession: T52028
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1081 <JOS>
A;Cross-references: EMBL:AF062485; PIDN:AAC29067.1

Query Match      62.0%; Score 3583; DB 2; Length 1081;
Best Local Similarity 61.0%; Pred. No. 2.7e-254;
Matches 676; Conservative 162; Mismatches 198; Indels 72; Gaps 15;

Qy      1 MDGADAA-LKSGRHGAGDVQCI CADGLGTLDGVFTACDVCRPPVCRPCYEHKRGQTQ 59
Db      17 INADENARIRSVQSLSGQTQICRDEITLVDSGPFVACNECAPVCRPCYEVERREGNQ 76
Qy      60 ACLOCKTKYKRRHSGPAIRGEGDDTDADGSDNPYPASGTE-DQOKIADRMRSWRMT 118
Db      77 ACPOCKTRFRRLKSGPRVGEDEDDIDLDLNEFEYGNNGIGFDQ---VSGMGSISRRNS 133
Qy      119 GGSNGVGHPRKYDSGEIGLSKYDSGEIIPRGYVPSVTNSQMSGEIPGASPDHH----- 169
Db      134 -----GFPQSD-----LDSAP-FGSQIPLLITYGDEDEVEI---SSDRHALIVPPLSG 175
Qy      170 -----MMSPGTNISRAPPYVNSHPNSRPSGSGIGNVAKRVDGWMKQDKG 219
Db      176 GHGNRVHPVSLSDPTVAHRRLMVPQKDLAVY-----GYGSVAKDRMBEWRKQNE- 227
Qy      220 AIPMTNGTSIAPSGRAATIDIDASTEYNMEDALLNDETROPSPKVPITASKINPYRMVI 279
Db      228 -----KLQVRHGGDDPDFEGDDADPFM-----MDEGRPLSMKPIPKSKINPYRMVI 276
Qy      280 VLRVLVLSILFHYRLTNVPRNAYPLWLLSVICETWFAISLWLDQFPKWFPTINRETYLDRL 339
Db      277 VLRVLVGLFPHYRLHPVNDAYALWLSVLCETWFAVSVLDQFPKWPYERITYLDRL 336
Qy      340 ALRYDRGSGPSQLAAVDIFVSTVDPLKEPPIVTANTVLSILAVDPVDKVCYVSDGAS 399
Db      337 SLRYEKSGKPSGLSPVDVFSVTDVPLKEPPIITANTVLSILAVDPVDKVCYVSDGAA 396
Qy      400 MLTFDALAEISFARKKVPVFKYKIDIPRAPEFFECOKIDYLDKQVPSFVKDRAMKRE 459
Db      397 MLTFEALSETAEFARKVVPFKCYICIEPRAPPEWYFCHMDYLNKNVHPAFVREBRAMKR 456
Qy      460 YEEFKIRNALVSALKVPSEGWTMQDGTWPNGNTRDHPGMIQVFLGSHGSLDTEGNEL 519

```

```

Db      457 YEEFKVKINALVATAQKVPEDGWTQDGTWPGNSVRDHPGMIQVFLGSDGVRDVNNEL 516
Qy      520 PRLVYYSREYRPGFHHKKAGAMNALVRYSAVLTNGOYMLNLDCHYINNSKAVREAMCF 579
Db      517 PRLVYYSREYRPGFHHKKAGAMNSLRVSGVLSNAPYLLNVDCDHYINNSKALREAMCF 576
Qy      580 LMDPNLGPQVYQVQFQRFQDIDRNDRYANRNTVFFDINLRGLDIOGQPVVVGTCVFNRR 639
Db      577 MMDPOSKKICYQVQFQRFQDIDRNDRYANRNTVFFDINLRGLDIOGQPVVVGTCVFNRR 636
Qy      640 TAIYGYEPPIKAKKPGFLAS-----LCGGKKKASKSKRSSDKKSKNKHVDSVSVFVN 692
Db      637 QALYGFDPAPKKKGPRTKNCMPKWCCLLFGSRKNRRAKTVAAADKKKNR--EASKQIHA 694
Qy      693 LEDIEEGVAGFDDKESVLMQMSLEKREGOSNAAFVASTILMEYGGVQPSSTPESLAKHA 752
Db      695 LENIEEGRHKVLNVEOSTEAMQMKQKQYGSQPVFVASARLENGGGMARNASPACLLKEA 754
Qy      753 IHVISCYEDKSEWGTETIGWYGSVTEDILTDFGKHARGWRSVYVCMPEKRFAPKGSAPINL 812
Db      755 IQVISRGYEDKTEWKEIGWYGSVTEDILTDFGKHARGWRSVYVCMPEKRFAPKGSAPINL 814
Qy      813 SDRNLQVLRWALGSVEILFGRHCPWYGYGRUKFLERPAYINTIYPLTSLPLLVYCIL 872
Db      815 SDRNLQVLRWALGSVEILFGRHCPWYGYGRUKFLERPAYINTIYPLTSLPLLVYCIL 874
Qy      873 PAICLLTGKEMPEISNLSIWFIALFSLFATGILEMRSGVGIDSEWRNEQFWITGGI 932
Db      875 PAICLLTGKEMPEISNLSIWFIALFSLFATGILEMRSGVGIDSEWRNEQFWITGGI 934
Qy      933 SAHLFAVFOGLLKVLAGIDNFTVTSKANDEEGFAELYFKFWITLIPPTILLINMGV 992
Db      935 SAHLFAVFOGLLKVLAGIDNFTVTSKANDEEGFAELYFKFWITLIPPTILLINMGV 993
Qy      993 VVAGTSYAINSGYSQSWGPFGLKFFAFVWVHLVFLKGLMGRQNRTPPTIVWAVLLAS 1052
Db      994 VVAGTSYAINSGYSQSWGPFGLKFFAFVWVHLVFLKGLMGRQNRTPPTIVWAVLLAS 1053
Qy      1053 IFSLLWVRVDPPTFTRLAGPNIOQTGNC 1080
Db      1054 IFSLLWVRVDPPTFTRLAGPNIOQTGNC 1080

RESULT 7
H84604
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84604
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vauken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84604
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1088 <STO>
A;Cross-references: GB:AE002093; NID:g4417271; PIDN:AAD20396.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g21770
A;Map position: 2

Query Match      61.9%; Score 3579; DB 2; Length 1088;
Best Local Similarity 61.0%; Pred. No. 5.3e-254;
Matches 677; Conservative 156; Mismatches 199; Indels 76; Gaps 20;

Qy      4 DADALKSGRHGAGDVQCI CADGLGTLDGVFTACDVCRPPVCRPCYEHKRGQTQACLO 63
Db      24 DTAIRSABELSGQTKICRDEITLVDSGPFVACNECAPVCRPCYEVERREGNQACPQ 83
Qy      64 CKTKYKRRHSGPAIRGEGDDTDADQ-----GSDFNYPASGTEQKQIADRMRSWRN 117

```


Db 84 CQTRYKRIKGSRVGEDEDD--DIDDLEHEFGMD---PEHVTE-----AALYVMRLN 132
QY 118 TG-GSGNVCHPKYDSEIGLSKYDSEIGIPRGVVPVSTNSQMSGEIPGASPDHMMSP---- 173
Db 133 TGRGTDEVSH-----LYSAS--PGSRVPLLTYYCDESDM--YSDRHALVPPST 177
QY 174 -TGNISRRAPF-----PYVNSPNPREFS-GSIGNVAKERVVDGWMQKQ-DKGA 220
Db 178 GLGNRVHVPFDTSPASITHRPWV-----PQDLTVYGYGSVANKORMEVWKKQIEKIQ 232
QY 221 IPMTNGTSTAPSEGRAATDIDASTEYNMEDAL-LNDETRQPLSRKVPPIASSKINPYMVI 279
Db 233 VVKNERNVNDGCGDGFVDBLD-----DPGLPMWDEGRQPLSRKLPPIRSSRINPYMLI 285
QY 280 VLRLVLSILFLHYRLTNVRNAYPLWLLSVLCEIWFALSWILDOPKMPPIRETYLDR 339
Db 286 FCRLLAILGLFFHYRLHPNDVAFGLWLTSLVCEIWFASWILDOPKMPPIERETYLDR 345
QY 340 ALRYDREGEPSSLAARVDFVSTVDPLKEPPIVTANTVLSILAVDYPVDKVSVCVSDDGAS 399
Db 346 SLRYEKECKPSLAPVDVFSVTDPLKEPPLITANTVLSILAVDYPVEKVACVSDDGA 405
QY 400 MLTTPALAESETFAKRWPFVKKYDIEPRAPPEFCQIDYLDKQVQPSFVKDRRAMKRE 459
Db 406 MLTPEALSVTAFAKRWPFVKKYDIEPRAPPEFCQIDYLDKQVQPSFVKDRRAMKRD 465
QY 460 YEEFKIRINALVSKALVYEEGWIMODGTPWPGNTRDHPGMIQVFLGHSGLDTEGNE 519
Db 466 YEEFKVKNALVSVQKPEDGWTQDGTWPGNVRDHPGMIQVFLGHSVCMDNGEL 525
QY 520 PRLVVSREKRGFQHHKAGAMNALVRVSAVLINGQYMLNDCDHYINNSKAVREAMCF 579
Db 526 PRLVVSREKRGFQHHKAGAMNSLIRVSAVLINAPYLLNVDCHYINNSKAIREAMCF 585
QY 580 LMDPNLGVQVCVQPPQDPDIDRNDRYANRNTVFFDINLRGLDGIQGVVYVGTGVNFR 639
Db 586 MMDPQSGKKICVQPPQDPDIDRNDRYANRNTVFFDINLRGLDGIQGVVYVGTGVNFR 645
QY 640 TAIYGEYPIKAKPGFLAS-----LCGGKKKASKKRSDDKKSKNKHVDSVPVN 692
Db 646 QALYGFDAKKQPPGRTCNWPKWCCCLCCGRRKKTKGVQDNQRRKKP--ETSKQIHA 702
QY 693 LEDIEGVEGAGFODKESVLMSSQMSLEKRFQSAAFVASTLMEYGGVPOSSTPESILLKEA 752
Db 703 LEHIEEGVQVTAENNSSET--AQLKLEKFGQSPVLVASTLLNGVPSNVNPPASLLRES 760
QY 753 IHVISCYGEDKSEWGTETIGWYGSVTEIDILTFQKMHARGWSVYCMKRPAPKGSAPINL 812
Db 761 IQVISCYGEETEWGKEIGWYGSVTEIDILTFQKMHARGWSVYCMKRAAPKGSAPINL 820
QY 813 SDRNLQVLRWALGSVEILFSRCHPLWYGVGRKFLERFAYINTTIYPLTSLPLAVYCIL 872
Db 821 SDRNLQVLRWALGSVEILFSRCHPLWYGVGRKFLERFAYINTTIYPLTSLPLAVYCSL 880
QY 873 PAICLLTGKFIPEISNLASIFALFISFATGILEMRWSGVGIDEMWRNEQFVWIGI 932
Db 881 PAICLLTGKFIPEISNAGILFLLMFNSIAVTGILEMQWKGIGIDDWRNEQFVWIGV 940
QY 933 SAHLFAVFGQLLKVLVAGDITNFTVTSKANDEBDFAEIYMKFWTLLIPPTTILINWVG 992
Db 941 SSHLFAVFGQLLKVLVAGDITNFTVTSKANDEBDFAEIYMKFWTLLIPPTTILINWVG 999
QY 993 VVAGTSYAINSGYQSGWPLFGKLFGLFAFVWVHLXPLKGLMGQRNRTTIVIVWAVLAS 1052
Db 1000 VIVGSDAINNGYDWSGWLFGELFALWVIVHLXPLKGLMGQRNRTTIVIVWAVLAS 1059
QY 1053 IFSLWVRVDPPTTLRAGNIQTGCINC 1080
Db 1060 ILTLLWVRVNPVPSK-DGPFVLEICGLDC 1086

Tl0797
cellulose synthase (EC 2.4.1.-) catalytic chain celA1 - upland cotton
C.Species: Gossypium hirsutum (upland cotton)
C.Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C.Accession: Tl0797
R.Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
A.Title: Higher plants contain homologs of the bacterial celA genes encoding the catalytic
A.Reference number: Z17152; MUID:97057296; PMID:8901635
A.Accession: Tl0797
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-974 <PE>
A.Cross-references: EMBL:U58283; NID:gl706955; PIDN:AAB37766.1; PID:gl706956
A.Experimental source: strain Acala SJ-2; fiber
C.Genetics:
A.Gene: celA1
C.Function:
A.Description: involved in the synthesis of cellulose
C.Keywords: glycosyltransferase; hexosyltransferase
Query Match 60.2%; Score 3477; DB 2; Length 974;
Best Local Similarity 61.5%; Pred. NO. 1.4e-246;
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;
QY 18 VQICACADGLGTTLDGDVFTACDVCRFPVPCPCYHERKEGTQACLOCKTKYKRHRGSPAI 77
Db 8 VCHTCGEHVGLNVNGEPFVACHCECNFPICKSCFFYDLKEGRKACLR-----GSPY- 58
QY 78 RGEEDDTDADGSDPNYPASGTEDQKQIADRMRSWRMTTGGSNVG-HPKYDSEIGL 136
Db 59 -----DENLDD-----VERATGQSTMAA-----HLNKSQDVGIHARHIS--SV 96
QY 137 SKYDSEIPRGVVPVSTNSQMSGEIPGASPDHMMSPGNTISRRAPFPVNVHSPSREF 196
Db 97 STLDS-----EM 103
QY 197 SGSGINAWKERVQGMKQDKGAIPMTNGTSTAPSEGRAATDIDASTEYNMEDALLNDE 256
Db 104 AEDNGNSTWKRVESEWKEKXKKKPAIT-----KVERAEIPEQOMEDPAPDA 154
QY 257 TRQPLSRKVPPIASSKINPYRMVILRLVVLVLSIFLHYRLTNVRNAYPLWLLSVI 316
Db 155 S-QPLSTIIPKSLAPYRTVIIMRLIILGLFPHYRTNPVDSAFGLWLTSLVCEI 213
QY 317 LSWILDQPKFPPIRETYLDRALRYDREGEPSSOLAVIDFVSTVDPLKEPPI 376
Db 214 FSWILDQPKFPVPIRETYLDRLSARYEREGERDELAADVDFVSTVDPLKEPPI 273
QY 377 LSILAVDYPVDKVSVCVSDDGASMLTFDALAETSEFARKWVPFVKYKDIETPRAPPEF 436
Db 274 LSILALDYPVDKVSVCYISDDGAMLTFFESLVETADFAKRWPFVCKKFSIEPRAPPEF 333
QY 437 KIDYLDKQVPSFVKDRRAMKREYEEFKIRINALVSKALKVPERGWINQDGTWP 496
Db 334 KIDYLDKQVPSFVKDRRAMKRDYEEFKIRINALVAKAQTDPDEGWTQDGTSW 393
QY 497 DHPGMIQVFLGHSGLDTEGNEPLRVVVSREKPGFOHHKAGAMNALVRVSAVL 556
Db 394 DHPGMIQVFLGYSRGARDIEGNELPRLVVSREKRPYQHHKAGAMNALVRVSAVL 453
QY 557 YMLNLDCDHYINNSKAVREAMCFMLMDPNLGPQVCVQPPQDFDIDRNDRYANRNTV 616
Db 454 FILNLDCDHYINNSKAVREAMCFMLMDPQVGRDVCVQPPQDFDIDRNDRYANRNTV 513
QY 617 INLRGLDGIQPVVYVGTGVNRTAIYEPPIKAKPGFLASLCGGKKKASKKRS 676
Db 514 VMKGLDGIQPVVYVGTGVNRTAIYEPPIKAKPGFLASLCGGKKKASKKRS 573
QY 677 KKSNKHVDSVPVFNLEIEEGVEGAGFDD-EKSVLMSQMSLEKRFQSGAAFAV 735
Db 574 LYRDAKREELDAIFNLREID-----NYDEYERMLISQTSFEKTFGLSVFI 627

Db 712 EAJAVISCWEDNTEWDRIGWYGVTDVVTGYRMHNRGWSVYCIITKRDAPRGTAPl 771
QY 811 NLSRLNQLRWALGSEVIELPSRCHPLWYGGRLKELERFAYINTIYPLTSLPLLYVC 870
Db 772 NLTDRLHQVLRWATGSVEIFFSKNAMF--ATRRLKELQRAVLNVGIGYPTIFLIVVC 829
QY 871 ILPAICLLTGKFMPEISNLASTWIALFSLFIPAT---GILEMRWSGVGIDEMWRNEQF 926
Db 830 FLPAICLFSKGFVQSJ---DIHFLSYLCITVTLTSLLELVKNSGIGLEWRNEQF 885
QY 927 WVIIGGSAHLFAVFGQLKLVLAGIDTNTVTYSKANDEGD--FAELMYFMKWTLLIPPTT 984
Db 886 WLIGGTSALHAAVVQGLKLVLAGIEISFTLTSKASGEDEDDIADLYIVKWTGLFIMPLT 945
QY 985 ILLINMVGWAGTSYAINSGYQSGWPLFGKLFAPFWIVHLYPLKGLMGRQRTPTIVI 1044
Db 946 IIVNLVAIVIGASRTIYSVPOWGLMGIGFFLSWLTHMYPPAKGLMGRGRKVTIYV 1005
QY 1045 VWAVLLASIFSLWVRVDP 1063
Db 1006 VNSGLVSIITVSLWITISP 1024

RESULT 15
C86446
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
R;Accession: C86446
R;Theologits, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-979 <S70>
A;Cross-references: GB:AE005172; MID:gl0801364; PIDN:AAG23436.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 37.4%; Score 2163.5; DB 2; Length 979;
Best Local Similarity 44.3%; Pred. No. 3.4e-150;
Matches 450; Conservative 173; Mismatches 245; Indels 147; Gaps 24;

QY 142 GEIPRGVPSVNTSQMGEIPGASPDHMHMSPTGNISRRAPFPYVNHSPNPSRFGSGIG 201
Db 4 GESPLRH-PRISHVNSGSDPGSSSDYKNY-----LVQ1PPTP--DNNPGA-----SLS 50

QY 202 NVAMKERVGCHKQDKGAIPMTNGTSTAPSEGR-----AATDIDASTEYNMEDALLNDET 257
Db 51 IVL-----LEIDSNQESVPSGSDIVSGSGKDNPDLDVIRNVGEEEDDTLLSKI 103

QY 258 RQPLSRKVPITASSKINPYRMVILRLVLSIFLHYLTNPVRNAYPLWLLSVICELWIFAL 317
Db 104 SYSTRVVKISPIIALYRIILVVRVSLALFLWRINPNKALWLLSVICELWIFAP 163

QY 318 SWILDQFPKWPINRETYDLRALRY-----DREGEPSQLAAVDI FVSTVDPLKEPPIVT 372
Db 164 SWLLDQIPKLPVFNHATDIEALKATFETPNPDNPTGKSDLPGLDVFVSTADAKEPEPLVT 223

QY 373 ANTVLSTILAVDYPVDKVSYSVDGASMLTDLALAESEFARKWVPFVKYDIEPRAPEF 432
Db 224 ANTILSTILSDVPYVEKLSVYISDDGSLVTFEIAEAAASFAKIWWPFCRKHKIEPRNPFS 283

QY 433 YFCQKIDYLDKQVQPSFVKDRRAMKREYEEFKIRNALV-----SK----- 473

Search completed: August 23, 2004, 01:09:14

Job time : 35 secs

Db 284 YFGLKRDYKDKVRHDFVRRERYVKRAYDEFKVRVNALPHSIRRRSDAFNSKEEIKALEK 343
QY 474 -----ALKVPEEGWIMODGTPWPG-----NNTRDHPGMIQVFL- 506
Db 344 WKHKVKKVEBDQIKPRPALVAPKATW-MSDGTHTWPGTWAIVSGPHHSRGDHASVIQVLLD 402
QY 507 -----GHSGGIDTEGNE--LPLRVVVSREKPGFQHHKAGAMNALVRVSALVT 553
Db 403 PGDEPVEGKGGRALDLEGVDIRLPLMIVYSREKPGYDHNKAGAMNALVRASIMS 462
QY 554 NGQYMLNLDCHYINNSKAVREAMCFMLDPNLGPQVCYQFQORFDGIDRNDRYANRNTY 613
Db 463 NGPFIINLDCDHVYNSRAFRDGCIFMMDHD-GDRVSVYQFQORFEGIDPSDRYANKTV 521
QY 614 FFDINRLDGIQGPVTVGTGCVFNRTAIYGYEPP-----IKAKPGFLASICGGKKKASKS 670
Db 522 FFDINRLDGIQGPVTVGTGCLFRRTALYGFNPPDVFVVEEFP--GSY-----FPLI 574
QY 671 KKRSSDKKSNKHVDSSVPVFNLEDIEEGVAGDFDEKSVLMSQMSLEKRFSGSAAFVA 730
Db 575 KKRSPATVAS-----EPEYTDDEDRFDIGL--IRKQFGSSSMLVN 613
QY 731 ST-LMEYGGVPOQST-----PESL-----LKEAIIHVISCGYEDKSEWGTET 770
Db 614 SVKVAPEGRPLATVHSSRLGRPPGSLTGSRKPLDPATVNEAVNVISCWYEDKTEWGFNV 673
QY 771 GWYGSVTEDLITGFKMHARGWSVYCMKPRPAKGSAPINLSDLNQLRWALGSEVIEL 830
Db 674 GWYGSVTEDVVTGFRMHKGRWSFYCVTBPDAFRGSAPINLTDRLHQVLRWATGSVEIF 733
QY 831 FSRHCPWLWYGVGRLEKELERFAYINTIYPLTSLPLLYVCILPAICLLTGKFMPEISNL 890
Db 734 FSRNNAIF--AGPKLKLQRIALNVGIIYPTSFILTYCFLPPLSLFSGHFVVEITGS 791
QY 891 ASIWFIALFLSIFATGILEMRWSGVGIDEMWRNEQFWVIGISAHLEFAVFGQLKVLVLAGI 950
Db 792 FLIYLLIITLSLCLGLAVLEVKWSGISLEWRNEQFWLIGTSAHLVAVLQGLKLVIAGV 851
QY 951 DTNPTVTSKA-----NDEGDPFAELYMFKWTLLIPPTIILINMVGWAGTSYAINSGYQ 1006
Db 852 EISFTLTSKSTGGDEDEDFADLYLFKWTALMIPPLTIILINIVAILFAVCRTVFSANP 911
QY 1007 SWGFLFKGLFFAFWVIVHLYPLKGLMGRQRTPTIVIVWAVLLASIFSLWVRV 1061
Db 912 QWSNLLGGTFFASWVLLHMYPPFAKGLMGRGKPTVVVWVWSGLIACLSLLYITI 966

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2004, 22:25:29 ; Search time 19 Seconds
(without alignments)
2959.777 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSGRHGAGDVQC.....VDPFTRLAGNITQCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	264	4.6	874	1	BCSA_SALTI
2	264	4.6	874	1	BCSA_SALTY
3	259.5	4.5	1596	1	ACG2 ACEXY
4	258	4.5	872	1	BCSA_ECO57
5	257	4.4	872	1	BCSA_ECOLI
6	248	4.3	739	1	BCSA_PSEFL
7	238.5	4.1	754	1	BCA1 ACEXY
8	238.5	4.1	1518	1	BCA4 ACEXY
9	238.5	4.1	1518	1	BCA5 ACEXY
10	237.5	4.1	756	1	BCA2 ACEXY
11	230	4.0	729	1	BCSA_XANAC
12	207	3.6	1550	1	ACG1 ACEXY
13	202.5	3.5	745	1	BCA3 ACEXY
14	119	2.1	1433	1	SUBF_BACSU
15	116.5	2.0	441	1	YDQM_ECOLI
16	112.5	1.9	322	1	NU1M_STRPU
17	112	1.9	1029	1	RIP3 RAT
18	111	1.9	414	1	YOAB_BACSU
19	110.5	1.9	1154	1	KDGD_MESAU
20	110	1.9	753	1	STL5 RAT
21	109.5	1.9	1024	1	RIP3 MOUSE
22	108.5	1.9	1262	1	TPP2 MOUSE
23	108	1.9	1167	1	CLAA_BACTU
24	108	1.9	1894	1	BGS2_SCHPO
25	107.5	1.9	775	1	ICF0_HSV11
26	107.5	1.9	920	1	DMF1_SCHPO
27	107.5	1.9	1671	1	DPOLG_PVRKO
28	107.5	1.9	3033	1	POLG_HCVJ6
29	107	1.9	730	1	STL5_HUMAN
30	107	1.9	1249	1	TPP2 RAT
31	106	1.8	357	1	G6PT_CANFA
32	106	1.8	1177	1	A11B_HUMAN
33	105.5	1.8	662	1	CYOB_BUCAI

34	105.5	1.8	886	1	YKR1_CABEL	P34307 caenorhabdi
35	105	1.8	357	1	G6PT_HUMAN	P35575 homo sapien
36	104.5	1.8	541	1	COX1_PODAN	P20681 podospira a
37	104.5	1.8	694	1	RP3A_HUMAN	Q9Y2J0 homo sapien
38	104.5	1.8	992	1	UVRA_MICLU	P13567 micrococcus
39	104	1.8	374	1	MTLD_BACHD	Q9K681 bacillus ha
40	104	1.8	495	1	AMPA_LEPIN	Q8F0Q1 leptospira
41	104	1.8	517	1	COX1_PARLI	P12700 paracentrot
42	104	1.8	1249	1	TPP2_HUMAN	P29144 homo sapien
43	103.5	1.8	530	1	CG2B_DROME	P20439 drosophila
44	103.5	1.8	622	1	COX1_BACSU	P24010 bacillus su
45	103.5	1.8	911	1	NIA_LYCES	P17570 lycopersico

ALIGNMENTS

RESULT 1
BCSA_SALTI
ID BCSA_SALTI STANDARD; PRT; 874 AA.
AC 082291.
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR STY4181 OR T3898.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RA MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RA MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
uridine 5'-diphosphate glucose to cellulose, which is produced as
an extracellular component for mechanical and chemical protection
at the onset of the stationary phase, when the cells exhibit
multicellular behavior (rdar morphotype). Co-expression of
cellulose and thin aggregate fimbriae leads to a hydrophobic
network with tightly packed cells embedded in a highly inert
matrix (By similarity).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
+ {(1,4)-beta-D-glucosyl}(N+1).
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
(c-di-GMP) (By similarity).
CC -!- PATHWAY: Bacterial cellulose biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Potential).
CC -!- DOMAIN: There are two conserved domains in the globular part of
the protein: the N-terminal domain (domain A) contains
the conserved DXD motif and is possibly involved in catalysis and

CC substrate binding. The C-terminal domain (domain B) contains the
 CC QXXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL627281; CAD08006.1; -
 DR EMBL; AE016847; AAO71373.1; -
 DR InterPro; IPR003919; Cell_synth A.
 DR InterPro; IPR001173; Glyco trans 2.
 DR Pfam; PF00535; Glycos transf 2; 1.
 DR PRINTS; PRO1439; CELLSYNTHASEA.
 KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
 KW Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.
 FT FT 441 501 CATALYTIC SUBDOMAIN B.
 FT TRANSMEM 30 50 POTENTIAL.
 FT TRANSMEM 151 171 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 525 545 POTENTIAL.
 FT TRANSMEM 547 567 POTENTIAL.
 FT TRANSMEM 592 612 POTENTIAL.
 FT TRANSMEM 634 654 POTENTIAL.
 FT TRANSMEM 668 688 POTENTIAL.
 FT TRANSMEM 833 853 POTENTIAL.
 FT ACT SITE 313 313 POTENTIAL.
 FT ACT SITE 457 457 POTENTIAL.
 FT BINDING 360 360 SUBSTRATE (POTENTIAL).
 FT BINDING 362 362 SUBSTRATE (POTENTIAL).
 SQ SEQUENCE 874 AA; 100020 MW; 3395D09CD0F51990 CRC64;

Query Match 4.6%; Score 264; DB 1; Length 874;
 Best Local Similarity 19.5%; Pred. No. 6.4e-11;
 Matches 159; Conservative 91; Mismatches 220; Indels 346; Gaps 30;

QY 261 LSRKVPASSKINRMVIVRLVLSIFLHRLT-----NPNVAYPLMLLSVICETW 314
 DB LVRRMPGRFSAL---MLIVLSITVSCRIVWRYTSTLNWDDPVLVCGLLILFAETYAW 244
 QY 315 FALSILWLDOPKWPINRRTYLDRLALRYDREGEPSQLAAVDIFVSTVDPLKEPPIVTAN 374
 DB 245 IVL--VLGYFQVWPLNRQP-----VPLPKMSQWPTVDIFVETYN---EDLVNVRN 291
 QY 375 TVLSILAVDPYDVKVSCYSDDGASMLTFDALAETSEFARKWPFVKKYDIEPRAPEYF 434
 DB 292 TYIASLGIDWPDKDNIWILDGGR----- 316
 QY 435 CQKIDYLDKQVQPSFVKDRRAMKREYEEFKIRINALVSKALKVPBEGWIMQDTPWGN 494
 DB 317 -----ESFR----- 320
 QY 495 TRDHPGMIQVFLGSHGLDTEGNELPRLVYVSEKRPQHKKAGAMNALVRVSAVLIN 554
 DB 321 -----OFARHV-----VHYIAR-----ATHEHAKAGINNALKHA-----K 352
 QY 555 GQYMLMLDCHYINNSKAVREAMCFMLDPNGLGQVCYVQPPQRF---DGDIDRN-----DRY 607
 DB 353 GERVAIEDCHVPTRSFLQMTGWFLE---KQLAMMQTPHHFFSPDPERNLGRPKT 408
 QY 608 ANRNTVFFDINRLGLDGIQGVVYVGVGVGNRTAIYGYEPTIKAKKPGFLASLCGGKKA 667
 DB 409 PNEGTLFYLQVQGNMDWDATFFCGSCAVIRR-----KP----- 442

QY 668 SKSKRRSDKKKSNKHVDSSVPFNLEDIEGVGAGFDEKSVLMSQMSLEKRFQSSAA 727
 DB 443 ----- 442
 QY 728 FVASTIMEYGVGPQSTPESLKEAHVISCYEDKSEWGTGIEWYGSVTDIILTGFKM 787
 DB 443 -----LDEIGGI-----AVE-----TVTDAHTSLRL 464
 QY 798 HARGHRSYVCMKPRAPFKGSAVINLSRLNQVRLWALGSVELLFSHCPLWYGGRLKF 847
 DB 465 HRRGTYSAYM--RIPQSLATGESUSAHQIRWARGMVQI-FRLDNPL---FGKGLKL 518
 QY 848 LERFAVINTIYPLTSLPLLVYCVILPAICLLTGKTFMPEISNLASIFIALFLSIPA-TG 906
 DB 519 AGRLCYLANMFHLSGISRIPLFLPTAPLAFLLHAYI-----IYAPALMALFVIVH 569
 QY 907 ILEMWMSGVIGDEWWRNEQFWVIGISAHLFVAVFQGLLKVLAGIDTN---FTVTSKANDE 963
 DB 570 MVHASLTNSKIQKVRH-SFW--SEIYETVLAWYIAPPTLVALINPHKGFNVTA----- 622
 QY 964 EGDPAELYMFKWTLLIPPTTLLIINMGVAVAGTSVAINSGYQSGPLGKLPFAFWVIV 1023
 DB 623 -GGLVEEKYVDW-VISRPYIFLVNLLGLGVAAGV-----WRYVYGP----- 661
 QY 1024 HLYPFLKGLMGQRNRTPTIVIVWAVLLASIFSLWV 1059
 DB 662 -----ENETLVIV-----SLVWV 675

RESULT 2

BCSA_SALTY
 ID BCSA_SALTY STANDARD; PRT; 874 AA.
 AC Q93IN2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
 GN BCSA OR STM3619.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14028;
 RA MEDLINE=21160181; PubMed=11260463;
 RX Zogaj X., Nimtz M., Rohde M., Bokkraz W., Roemling U.;
 RT "The multicellular morphotypes of Salmonella typhimurium and
 RT Escherichia coli produce cellulose as the second component of the
 RT extracellular matrix.";
 RL Mol. Microbiol. 39:1452-1463(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE=21927388; PubMed=11929533;
 RA Solano C., Garcia B., Valle J., Berasain C., Ghigo J.-M., Gamazo C.,
 RA Lasa I.;
 RT "Genetic analysis of Salmonella enteritidis biofilm formation:
 RT critical role for cellulose.";
 RL Mol. Microbiol. 43:793-808(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea I., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes

uridine 5'-diphosphate glucose to cellulose, which is produced as an extracellular component for mechanical and chemical protection at the onset of the stationary phase, when the cells exhibit multicellular behavior (rdar morphotype). Co-expression of cellulose and thin aggregative fimbriae leads to a hydrophobic network with tightly packed cells embedded in a highly inert matrix.

-|- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP + {(1,4)-beta-D-glucosyl}(N+1).

-|- COFACTOR: Magnesium (By similarity).

-|- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid (c-di-GMP) (By similarity).

-|- PATHWAY: Bacterial cellulose biosynthesis.

-|- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).

-|- DOMAIN: There are two conserved domains in the globular part of the protein: the N-terminal domain (domain A) contains the conserved DXD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the QXXRW motif and is present only in processive glycosyl transferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.

-|- SIMILARITY: Belongs to the glycosyltransferase family 2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ315770; CAC44015.1; -

EMBL; AE008867; AAL22479.1; -

EMBL; AJ315148; CAC86199.1; -

StyGene; SG27222; bcsA.

InterPro; IPR003919; Cell_synth_A.

InterPro; IPR005150; Cellulose_synth.

InterPro; IPR001173; Glyco_trans_2.

Pfam; PF03552; Cellulose synt; 1.

Pfam; PF00535; Glycos trans 2; 1.

PRINTS; PR01439; CELLSTHASEA.

Cellulose biosynthesis; Transferase; Glycosyltransferase; Transmembrane; Inner membrane; Complete proteome.

DOMAIN 271 364 CATALYTIC SUBDOMAIN A.

DOMAIN 441 501 CATALYTIC SUBDOMAIN B.

TRANSMEM 30 50 POTENTIAL.

TRANSMEM 151 171 POTENTIAL.

TRANSMEM 173 193 POTENTIAL.

TRANSMEM 230 250 POTENTIAL.

TRANSMEM 525 545 POTENTIAL.

TRANSMEM 547 567 POTENTIAL.

TRANSMEM 592 612 POTENTIAL.

TRANSMEM 634 654 POTENTIAL.

TRANSMEM 668 688 POTENTIAL.

TRANSMEM 833 853 POTENTIAL.

ACT SITE 313 313 POTENTIAL.

ACT SITE 457 457 POTENTIAL.

BINDING 360 360 SUBSTRATE (POTENTIAL).

BINDING 362 362 SUBSTRATE (POTENTIAL).

SEQUENCE 874 AA; 100043 MW; 4C9421B58606310A CRC64;

Query Match 4.6%; Score 264; DB 1; Length 874;

Best Local Similarity 19.5%; Pred. No. 6.4e-11;

Matches 159; Conservative 91; Mismatches 220; Indels 346; Gaps 30;

261 LSRKVPASSKINPYRMVILRLVLSIFLHYLT-----NPVRNAYPLWLSVCEI 314

189 LVRRMPGRFSAL----MLIVSLTVSCRYIWMRYTSTLNWDPPSVLVCGLILLFAET 244

315 FALSWILDQFKWPFIPNRETLDRLALRYDREGPSQLAAVDIFVSTVDPLKEPIVTAN 374

245 IVL--VLGVQVVMPLNRQ-----VPLPKEMSQWPTVDFVPTYN---EDLNVRKN 291

375 TVLSILAVDPVDKVCYSYDDGASMLTFDALAETSEFAKWKVPFVKYKIDIEPRAPFYF 434

292 TIYASLIDGDPDKLNIWILDG-----K 314

435 CQKIDVLKDKVQPSFVKDRRAMKREVEEFKIRINALVSVKALVPEEGWIMQDGTWPWGN 494

315 -----GRE 317

495 TRDPHGMQVQLGHSGGLDTGEGNELPLRVVYSREKRPFGQHHKAGAMNALVRVSAVLTN 554

318 SPRH-----FAHVG-----VHYIARTT-----HEHAKAGINNALKHA---K 352

555 GQVWMLNDCDHYINNSKANREACFLMDPNLGPQVCVQFPQRF---DGIDRN---DRY 607

353 GEFAVAFDCHVPTRSFLQMTGWFKE-----KOLAMQTPPHFPSPDPERNLGRFRKT 408

608 ANRNTVFFDINLRGLDGIQGVVVGTCVFNRTAIYGYEPPKAKKPGFLASLGGKKKA 667

409 PNEGTLFYGLVQDNDWDATFCGSCAVIR-----KP-----442

668 SKSKRSSDKKXKHKVDSVPVFNLEDEBEGVEGAGFDDEKSVLMSQMSLEKRFQGSAA 727

443 -----442

728 FVASTLMYEGVQPSSTPESLLKEAHIVISGVEDKSEWTEIGWIYGSVTEIDLTGFKM 787

443 -----LDEIGGI-----AVE-----TVTEDAHTSLRL 464

788 HARGWRSVCMKRPAPKGSAPINLSRLNQVLRWALGVSVEILFSRHCPLWYGGYGLKLF 847

465 HRRGYTSAYM--RIPQAAGLATESLSAHIGQIRNARGMWQI--PRLDNPL---FGKGLK 518

848 LERFAYINTTIYPLTSLPLVYICLPAICLLTGKFMPEISNLASIWFIALFUSIFA-TG 906

519 AQRLCYLNAMFHLGSGIPRLIFLTAPLAFLLHAYI-----IYAPALMIALFVIPH 569

907 ILEWRWSGVGIDWRNEQFWIGGISAHLFANFQGLLKVLGIDTN---FTVTSKANDE 963

570 MVHASLTNSKIQGYRH-SFW--SEIYETVLAWIAPPTLVALINPHKGFNVTAK----622

964 EGDFAELYMFKWTLLIPPTILINMVGVVAGTSYAINSGYQSWGPLFGKLPFAFWIV 1023

623 -GGVVEEKYDVM-VISRPYIFLVLLNLLGVAAGV-----WRYYYGE-----661

1024 HLYPFLKGLMGRQNRTPTIIVVMAVLASIFSLMW 1059

662 -----ENETLTIV-----SLWVV 675

RESULT 3

ACSD ACXY ID ACS2 ACXY STANDARD; PRT; 1596 AA.

AC Q59167; 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Cellulose synthase 2 [includes: Cellulose synthase catalytic subunit (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose synthase 2 regulatory domain)].

GN ACSAII.

OS Acetobacter xylinus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;

OC Acetobacteraceae; Gluconacetobacter.

OX NCBI_TaxID=28448;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 23769;

RX MEDLINE=95394846; PubMed=7665515;

RA Saxena I.M., Brown R.M. Jr.;

RT "Identification of a second cellulose synthase gene (acsAII) in Acetobacter xylinum.";

QY 1019 FWIVHLYPFLKMGQRNRTETIVW-----AVLLASIFSLWV 1059
 Db 632 DWISRPYIEVLL-----NLGVAVGLWRYFYGPTEMLTVVWSWVW 675

RESULT 6

BCSA_PSEFL STANDARD; PRT; 739 AA.
 AC PS8931; Q8R8Z1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
 GN BCSA OR WSSB

OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=294;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=SBW25;

RX MEDLINE=22013850; PubMed=12019221;

RA Spiers A.J., Kahn S.G., Bohannon J., Travisano M., Rainey P.B.;
 RT "Adaptive divergence in experimental populations of pseudomonas
 fluorescens. I. Genetic and phenotypic bases of wrinkly spreader
 fitness";

RL Genetics 161:33-46(2002).

CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as
 CC an extracellular component responsible for the structural
 CC integrity and rigidity of self-supporting mats characteristic of
 CC the "wrinkly spreader" phenotype.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP
 CC + ((1,4)-beta-D-glucosyl)(N+1).

CC -1- COFACTOR: Magnesium (By similarity).

CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
 CC (c-di-GMP) (By similarity).

CC -1- PATHWAY: Bacterial cellulose biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).

CC -1- DOMAIN: There are two conserved domains in the globular part of
 CC the protein: the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC QXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AY074776; AAL71842.1; -

DR InterPro; IPR003919; Cell_synth A.

DR InterPro; IPR001173; Glyco_trans 2.

DR Pfam; PF00535; Glycos_transf_2; 1.

DR PRINTS; PR01439; CELL5NHASEA.

KW Cellulose biosynthesis; Transferase; Glycosyltransferase;

Transmembrane; Inner membrane.

FT DOMAIN 157 250 CATALYTIC SUBDOMAIN A.

FT DOMAIN 327 387 CATALYTIC SUBDOMAIN B.

FT TRANSMEM 36 55 POTENTIAL.

FT TRANSMEM 59 76 POTENTIAL.

FT TRANSMEM 83 101 POTENTIAL.

FT TRANSMEM 116 138 POTENTIAL.

FT TRANSMEM 417 436 POTENTIAL.

FT TRANSMEM 440 462 POTENTIAL.

FT TRANSMEM 524 546 POTENTIAL.
 FT TRANSMEM 551 573 POTENTIAL.
 FT ACT_SITE 199 199 POTENTIAL.
 FT ACT_SITE 343 343 POTENTIAL.
 FT BINDING 246 246 SUBSTRATE (POTENTIAL).
 FT BINDING 248 248 SUBSTRATE (POTENTIAL).
 SQ SEQUENCE 739 AA; 82165 MW; 2B962EA3854B23BB CRC64;

Query Match 4.3%; Score 248; DB 1; Length 739;
 Best Local Similarity 19.4%; Pred. No. 7.1e-10;
 Matches 155; Conservative 85; Mismatches 212; Indels 348; Gaps 34;

QY 263 RKVPIASSKINPYRMVILRLVLSI-----FLHYRLTNPNVRYAPLWL-----LSVICE 312
 Db 77 RKIP-----GRLLALVLSLVASLRVFWRLTSTL--GFETWDMFQYGLVAA 125
 QY 313 INFAL-----SWLDQFPKWFPIINREYLDRLALRYDREGPEPSQAALAVDFVSTV 362
 Db 126 EFVALIVLIFGYVQTAWPLRRTPWL-----KTEPEWPTVDVFITY 168
 QY 363 DPLKEPPIVTANTVLSILANDYPVDKVCYVDDGASMLTFDALAETSEFARKWVPFVK 422
 Db 169 N---EALSIVKLTIFAQAQMDPKDKLRVHVLDDG----- 200
 QY 423 YDIEPRAPEF-YFCOKIDYLDKVPQSFVKDRAMKREYEFEFKIRINALVSKALKVPEEG 481
 Db 201 -----RDDDFRFRCKV----- 212
 QY 482 WIMQDGTPEFNNTRDHPGMIQVPLGHSGGLDTEGNELPRLVYVSREKRGFQHHKAGA 541
 Db 213 -----GVN-----YIRDN-----NPHAKAGN 229
 QY 542 MNALVRVSAVLTVNGOYMLNLDCHYINNSKAVREAMCFMDNPLGPOVCVYQVQRF----- 598
 Db 230 LINEALKV-----TDGEYIALEADADHVPTRSLFQVSLGWLKDPKL-----AMLQTPHFFSP 281
 QY 599 DGDIDRN-DRY---ANRNTVFDDNLRLGDLGIQGVVYVGTGCVFNRTAIYGYEPPKAKKP 654
 Db 282 DPEKNUDTFRVNPNEGELFYGLVQDNDLWNATFFCGSCAVIRR-----EP----- 328
 QY 655 GFLASLGGKKKASKSKRSDDKKSNKHVDSSVPVFENLEDIEBEGVEGAFDDEKSVLMS 714
 Db 329 ----- 328
 QY 715 QMSLEKRGFQSAAFVASTLMYEGYVPOSSTPESLLKEAIIHVISCGYEDKSEWGTEIGMIY 774
 Db 329 -----LLEIGV-----AVE----- 338
 QY 775 GSVTDELITGFMHARGMRSVYCMKPRPAFKGSAPINLSDRLNOVLWALGSVILFSRH 834
 Db 339 -TVTEDAHTALKUNRLGNTAYL--AIPOAGLATESLSRHINORIRWARGMAQI-PRTD 394
 QY 835 CPLWYGYGGRKLFLERFAYINTTIYPLTSLPLVYCYLPAICLLTGKFIEMPEISLASIW 894
 Db 395 NPL---LGKGLKWCQRICYANAMQHFFYGLPRLVFLTAFLAYLIFG---AEIFHASALM 447
 QY 895 FIALFL-SIFATGILENRNSGVGIDENWRNEQF-----WVIGGISAHILFAVQGLLKVLA 948
 Db 448 IVAYVPLPHLVHSSLTNSRIQGRFRHSFW-NEVYETVLAWYI-----LPPVLVALVNPKA 500
 QY 949 GIDTNFTVTSKANDEEDGFARLYMFKWTTLLIPPTTILINMVGWAGTSYAINSGYQSW 1008
 Db 501 G---GFNVTDK-----GGIDKQFFDW-KLARPYVLVLANVLGLGFGIHLI-----W 545

QY 1009 GPLFGKLFPAF---WVIVHL 1025

Db 546 GDASTAVTVAINLTWTLNLYN 565

RESULT 7

BCAL_ACEXY

ID -BCAL_ACEXY STANDARD; PRT; 754 AA.

AC P19449;

DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
 GN BCSA.
 OS Acetobacter xylinus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Acetobacteraceae; Gluonacetobacter.
 OX NCBI_TaxID=28448;
 RN [1]__
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-42 AND 196-206.
 RC STRAIN=1306-3;
 RX MEDLINE=91045951; PubMed=2146681;
 RA Wong H.C., Fear A.L., Calhoun R.D., Eichinger G.H., Mayer R.,
 RA Amikam D., Ben-Ziman M., Gelfand D.H., Meade J.H., Emerick A.W.,
 RA Bruner R., Ben-Bassat A., Tal R.;
 RT "Genetic organization of the cellulose synthase operon in Acetobacter
 RT xylinum";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8130-8134(1990).
 RN [2]
 RP ENZYME REGULATION.
 RC STRAIN=1306-3;
 RX MEDLINE=21194671; PubMed=11297407;
 RA Chang A.L., Tuckerman J.R., Gonzalez G., Mayer R., Weinhouse H.,
 RA Volman G., Amikam D., Ben-Ziman M., Galles-Gonzalez M.-A.;
 RT "Phosphodiesterase A1, a regulator of cellulose synthesis in
 RT Acetobacter xylinum, is a heme-based sensor";
 RL Biochemistry 40:3420-3426(2001).
 CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
 CC uridine 5'-diphosphate glucose to cellulose in a processive way.
 CC The thick cellulosic mats generated by this enzyme probably
 CC provide a specialized protective environment to the bacterium.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
 CC + {(1,4)-beta-D-glucosyl}(N+1).
 CC -!- COFACTOR: Magnesium.
 CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
 CC (C-di-GMP).
 CC -!- PATHWAY: Bacterial cellulose biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- INDUCTION: Cellulose is produced at a linear rate with respect to
 CC cell growth when O(2) is present.
 CC -!- DOMAIN: There are two conserved domains in the globular part of
 CC the catalytic subunit: the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC OXXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M37202; AA21884.1; -.
 DR PIR; A43735; A43735.
 DR InterPro; IPR003919; Cell synth A.
 DR InterPro; IPR001173; Glyco trans 2.
 DR Pfam; PF00535; Glycos transf 2; 1.
 DR PRINTS; PR01439; CELLSYNTHASEA.
 DR Cellulose biosynthesis; transferase; Glycosyltransferase;
 KW Transmembrane; Inner membrane.
 FT DOMAIN 147 240
 FT DOMAIN 317 377 CATALYTIC SUBDOMAIN A.
 FT TRANSMEM 26 46 CATALYTIC SUBDOMAIN B.
 FT TRANSMEM 47 67 POTENTIAL.
 FT TRANSMEM 108 128 POTENTIAL.

FT	TRANSMEM	167	187	POTENTIAL.
FT	TRANSMEM	407	427	POTENTIAL.
FT	TRANSMEM	430	450	POTENTIAL.
FT	TRANSMEM	468	488	POTENTIAL.
FT	TRANSMEM	516	536	POTENTIAL.
FT	TRANSMEM	549	569	POTENTIAL.
FT	TRANSMEM	189	189	POTENTIAL.
FT	ACT SITE	333	333	POTENTIAL.
FT	ACT SITE	333	333	POTENTIAL.
FT	BINDING	236	236	SUBSTRATE (POTENTIAL).
FT	BINDING	238	238	SUBSTRATE (POTENTIAL).
SQ	SEQUENCE	754 AA;	84442 MW;	8DSFCIF62E2C068 CRC64;

Query Match 4.1%; Score 238.5; DB 1; Length 754;
 Best Local Similarity 20.1%; Pred. No. 3.6e-09;
 Matches 163; Conservative 82; Mismatches 232; Indels 335; Gaps 32;

QY	279	IVLRVLSISFLHYRLTNPVRNAYPIWLLSVICETWFALSILDO---PPKWFP-----	329
DB	53	LIVLVLCVLFVFIIVGRGSRRTQIFLEVLISALVSLRY-LTWRLTETLDFDTWIOGGLGVT	111
QY	330	-INRETY-LDLRLALRYDREGPSQLA-----AVDIFVSTVDPLKEPPIVTANT	375
DB	112	LIMAEYALYMLFLSYFQTIQPLHRAPLPDPNDVDDWPTVDIFITYD---EQLSIVRLT	168
QY	376	VLSILAVDVPVDKVCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPRAPEFYFC	435
DB	169	VLGALGIDWPPDKVNVILDDG-----	190
QY	436	QKIDYLDKVPSPFVKDRRAMKREVEEPKIRINALVSKALKVPBERGWMQDTPPGNNT	495
DB	191	-----VRP-----EPEFAKDCGAL-----	205
QY	496	RDPGMIQVFLGHSGGLTEGNEPLRLVYVSRKEPGFHKKKAGAMNALVRVSAVLING	555
DB	206	-----YIGRVDSS-----HAKAGNLNHAIK-----RUSG	229
QY	556	QYMLMLDCHYNNKSKAVREACFLMDNLPQVCYQVPPQPFQDFGIDRNDRYANENTVFF	615
DB	230	DYILLDCDH-IPTRAFQLIANGWVADR---KIALMQTPHHFYSPD-----	272
QY	616	DINRLGLDGIQGPVYVGTGVNRTAIYGYBPPIKAKPGFLASLCGGKKKASKSKRS	675
DB	273	-----PFQRNLAVGYRTP-----	285
QY	676	DKKSKNKHVDSSVPVFNLEDEEGVEGAGFDEKSVLMSQMSLEKRFQSOAFAVASTIME	735
DB	286	-----PEGNL---FYGVIQDND-----FWDATFFCGSCAI-313	
QY	736	YGGVPQSSTPESLLKEAIIHVISCGYEDKSEWGTETIGWYGVSTEDILTGFKNHARGWSV	795
DB	314	-----LRRAIESIG-----GPAVETVTEDAHTALRMQRGWSTA	348
QY	796	YCMKPRPAFKSAPINLSDRLNQLVRLWALGSVEILFSRHCLWYGYGRLLKFLERFAYIN	855
DB	349	YL--RIPVASGLATERLTHIGQRMWARGMIQI--FRVDNPM---LGGLGLQGLCYLS	402
QY	856	TTIYPLTSLPLVLCILPAICLLTGKFTMPISNLASIFWIALFSLFATGILEMWS--	913
DB	403	AMTSEFFFAIPRVIFLASPLAFPFQNTII-----AASPLVAYAIPIPMFHGIA	451
QY	914	-GVGIDEWNRNEQFW-VIGGISAHLFAVFQGLLKLVLAGIDNTFTVTSKA---NDEGDFA	968
DB	452	TAAKVKNKGR-YSEWSEVYETTMALFLVRVTIITLMPFSKGFENVTEKGVLEEEFDLG	510
QY	969	ELYMEKWTLLIPTTILLINWGVVAGTSVAINSGYOSWGPFLFG--KLFAFWIVVHLY	1026
DB	511	ATY-----PNIIF-----AGIMTLGLLGLFELTFHF-----	537
QY	1027	PFLKGLMGRQNTPTPTIVWA-----VLLAST	1053
DB	538	---NQLAGIAKRAVILNLCINWAMISILILLAAI	566

Db 247 QISGMWVSDS---NIALQTPHFYSPD----- 272
QY 634 GCVENRTAIVGYEPIVAKKPGFLASICGGKKASKSKKSKNDVSSVPPVNL 693
Db 273 --PQORLAVGYRTP-----PBGNL 290
QY 694 EDIEGVEGAGFDDDEKSVLMSQMSLEKRFGOSAAFVASTLMEYGGVQSSPTPESLKKEAI 753
Db 291 ---FYGIQDND-----FWDATFCGSCAI-----LRRKAI 319
QY 754 HVISCGYEDKSEWCTEIGWIYGSYTEDILTGFKMHARGWSVYCMRPKPAFKGSAPINLS 813
Db 320 EEIG-----GPAETVTEDAHTALRMQRKWSYAYL--RIPLASGLATERLI 364
QY 814 DRLNQVLRWALGSVEILFSRHCPWYGVYGRHLKELERPAYINTIYPLTSLPLLYCILP 873
Db 365 THIGQRMRWARGMTQI-FRVDNPM---LGSGKLGQRCLVLSAMTSFFFAIPRVIFLASP 420
QY 874 AICLLTGKFMPEISNLASTWFIALFLSIPATGILEMRWSGVGIDWNRNEQFW--VIGG 931
Db 421 LALFFSQNIIASPLAVGVAIAPHMPSHIAT-----AAKNKQWR-YSFWSVET 471
QY 932 ISAHFAVFGQLKVLADITNFTVTSKA---NDEEGDFAELYMKFTTLLIPTTILII 988
Db 472 VMA-LFLVRVITVMTLFPSSKGVNTEKGVLEREEFDLTATY-----PNIIFAIL 521
QY 989 NMGVVAVG-----TSYAINSGVQSGWGLFGKLFPAFWIVVHLYPLKGL-M 1033
Db 522 MALGLRGLYALIFQHLDIISERAYALN-----CIWSVISIILMAVISV 566
QY 1034 GRQNR 1038
Db 567 GRETK 571

RESULT 14

SUBF_BACSU
ID SUBF_BACSU STANDARD; PRT; 1433 AA.
AC P16397;
DT 01-FEB-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bacillopeptidase F precursor (EC 3.4.21.-) (Esterase) (RP-I protease)
DE (90 kDa serine proteinase).
GN BPR OR BPR OR BSU15300.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 195-222.
RX MEDLINE=90170864; PubMed=2104512;
RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,
RA Pero J.;
RT "Bacillopeptidase F of Bacillus subtilis: purification of the protein
RT and cloning of the gene.";
RL J. Bacteriol. 172:1470-1477 (1990).
RN [2]
RP REVISIONS.
RC STRAIN=168;
RX MEDLINE=90368623; PubMed=2118514;
RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,
RA Pero J.;
RL J. Bacteriol. 172:5520-5521 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90216713; PubMed=2108961;
RA Wu X.-C., Nathoo S., Pang A.S.-H., Carne T., Wang S.-L.;
RT "Cloning, genetic organization, and characterization of a structural
RT gene encoding bacillopeptidase F from Bacillus subtilis.";
RL J. Biol. Chem. 265:6845-6850 (1990).
RN [4]
RP SEQUENCE FROM N.A.

RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara K., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prasecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256 (1997).
RN [5]
RP SEQUENCE OF 1-211 FROM N.A.
RX MEDLINE=89008108; PubMed=3139638;
RA Beall B., Lowe M., Lutkenhaus J.;
RT "Cloning and characterization of Bacillus subtilis homologs of
RT Escherichia coli cell division genes ftsZ and ftsA.";
RL J. Bacteriol. 170:4855-4864 (1988).
RN [6]
RP SEQUENCE OF 1410-1433 FROM N.A.
RX STRAIN=168 / Marburg;
RC MEDLINE=90174995; PubMed=2106671;
RA Masuda E.S., Anaguchi H., Sato T., Takeuchi M., Kobayashi Y.;
RT "Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus
RT subtilis.";
RL Nucleic Acids Res. 18:657-657 (1990).
RN [7]
RP SEQUENCE OF 195-219.
RC STRAIN=Natto 16;
RA Kato T., Yamagata Y., Arai T., Ichishima E.;
RT "Purification of a new extracellular 90-kDa serine proteinase with
RT isoelectric point of 3.9 from Bacillus subtilis (natto) and
RT elucidation of its distinct mode of action.";
RL Biosci. Biotechnol. Biochem. 56:1166-1168 (1992).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M29035; AAA62679.1; -;
DR EMBL; J05400; AAA83362.1; -;
DR EMBL; Z99111; CAB13403.1; -;
DR EMBL; M22630; AAA22458.1; -;
DR EMBL; X17344; CAA35224.1; -;
DR PIR; A36734; A36734.

DR HSP; P00782; 2SBT.
DR MEROPS; S08.017; -.
DR SubtilList; BG10233; bpr.
DR InterPro; IPR008757; Peptidase M6.
DR InterPro; IPR000209; Peptidase S8.
DR InterPro; IPR009020; Protease inhib.
DR Pfam; PF05447; Peptidase M6; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE H18; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
DR KW Hydrolyase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 30
FT PROPEP 31 194
FT CHAIN 195 755
FT PROPEP 756 1433
FT ACT_SITE 227 227
FT ACT_SITE 274 274
FT ACT_SITE 452 452
FT CONFLICT 219 219
FT CONFLICT 393 393
FT CONFLICT 829 834
FT CONFLICT 836 841
FT CONFLICT 844 852
FT CONFLICT 853 1433
SQ SEQUENCE 1433 AA; 154577 MW; 98DF6846897807C9 CRC64;

Query Match 2.1%; Score 119; DB 1; Length 1433;
Best Local Similarity 19.8%; Pred. No. 4.7; Mismatches 96; Indels 334; Gaps 49;
Matches 184; Conservative 96;

QY 13 HG---AGDVQCICADGLGTLTLDGVFTACDVCRFPVCRPCYBHER-----54
DB 503 HGLVNAFAVSAVTDLGLKA-EGQSVSGDDDEPPV---YQHKVTEAYEGSLPLTIL 557
QY 55 KEGTQACLCQCKTKYKRRHG-----SPAIRBEGDDTDADGSDFNYPASGTEDQKQITA 108
DB 558 AEDNVSVTSVKLSYKLDQGEWTEITAKRISGDHLKGT-----YQAEIPDIKGTKLS 608
QY 109 DMRSWRMNTGSGNVGHKPYD-----SCEIGLSKYDSGEIPRGVVPVTSNQSOMGEIPG 163
DB 609 YK---WMTHDFGHVSSVDYDVTVKPSTAGY-KODFETAPGGVWASGNTNNWEGVPS 664
QY 164 ASPDH-----HMSPTGNIRAFPPVYNHSPNPSREFSGSI-----200
DB 665 TGPNTAASGEKVGYNLTGNVANSANMLV--MPPIKAPDSGLFLQPKSWHLEDDPDY 722
QY 201 -----GNVANK-----BRVD--GWKMKQDKGAIPMTNGTSTIAPSEGRA 236
DB 723 GYFVFLPEGEKNWEQAGVYNGKTSWTDEIDLSAYKQNIQVMFNLSDESIA-KEGWY 781
QY 237 ATDI---DASTEYNMEDALLNDETQPLSRKVPYIASSKINPYRMVILVLSIFLHYR 293
DB 782 IDVVVLSKSAKTVKKNKLGVEKPSGQKKKPVNPKAKPSANTAVKH-----QNK 833
QY 294 LTNPVNRNAYPLWLLSVICIEWPAFSLSWILDQPFKFPINRETLDRLALRY-----DREGES 350
DB 834 ALQP--QVPLKAQSVVE-----TGKSTYSDOSTGYTLKHAGDYT 874
QY 351 QLAADVIFVSTVD--PLKEPPIVTVANTVLSIL-----ADVVPVDKVSCVSDSD 396
DB 875 LMAEAYGQSKTKVSLKTDQTTQANFTLEEMKGTGLKGTINKTTPGVPVTCASVYVVED 934
QY 397 GA--SMLTFDA--LAETSEFARKWVFFVKVYDIEPRAPEFYFCQKIDYLDKQVSPFK 451
DB 935 AAVEPAMNDKGEYMLEAYEGA-----YTIKVAAPGYI-----967
QY 452 DRAMKREYEFKIRINALVSK--ALK--VPEEGWIMQD-----GTPWPGNN 494
DB 968 -----SDEFSVELKGDVTKETALKPVPVGPGEIAYDDGAENANSYFAAGNGWAVKM 1019
QY 495 T-----RDHPGMLQVFLHSGGL-----DTE-----GNELPLVY--VSREKRFQFHKKA 539

DB 1020 TLADGKD-KGML-----TGGLFREWDTEFPDGGTEFKFVEYDATGKGAPG---KKIA 1069
QY 540 GAMNALVRVSAVLITNGQYM-----LNLDCDHVYNNKAVREAMCFLMDNLPQVQCY 591
DB 1070 GEFNA-----EALRNGEWTKVDLSKSGIMVDKDFYL-----VY 1102
QY 592 VQF-PQRFD---GIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGBE- 646
DB 1103 IQSKPDPPSPGLAMDETQNSGRNWQYIDGKWQPGDKADGNMYI-----RALVDYEA 1154
QY 647 --PPIKA-----KKPGLASLCGGKKKASKSK-----671
DB 1155 AVPEITSPDKSYTNKDSVTVKGNASPGTTHIYNGEKEAGETKAAADGTFHAGIILNG 1214
QY 672 --KRSSDKKSNKHVDSSVPV-----PNLEDIBEG-----VEGAGFDDE-K 709
DB 1215 ENELTATATSTONGTTDASPTITVLDQKPELTLDNPKDGGTKNKETLTVRGAVSDDNLU 1274
QY 710 SVLMSQMSLEKRFQSGSAFVASTLMEYG 737
DB 1275 DV---KVGKKATVADGYSARILLENG 1299

RESULT 15
YCDQ_ECOLI
ID YCDQ_ECOLI STANDARD; PRT; 441 AA.
AC P75905;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ycdQ.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO Y.PESTIS HMS LOCUS PROTEIN HMSR AND TO
S.BPIDERMIDIS ICAM.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000204; AAC74107.1; -.


```
Db 61 CPQCKTKYKRHKGSFAIRGEEDGTDADDVSDYNYPASGSADQKQKIADRMRSWRMNAAG 120
Qy 121 SGNVCHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db 121 GGDVGRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMSPTGNIGKR 180
Qy 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTIAPSEGRAATI 240
Db 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTIAPSEGRVGD 240
Qy 241 DASTYNNMEDALLNDETQPLSRKVPFIASSKINPYRMVIVLRLVLSIFLHYRLINPVN 300
Db 241 DASTYNNMEDALLNDETQPLSRKVPFIASSKINPYRMVIVLRLVLSIFLHYRLINPVN 300
Qy 301 AYLPLMILLSVICIWFALSILANDYVPDKVSCVSDDGASMLTFDALAETSEFARKWVPF 360
Db 301 AYLPLMILLSVICIWFALSILANDYVPDKVSCVSDDGASMLTFDALAETSEFARKWVP 360
Qy 361 TVDPLKEPPIVNTANTVLSILAVDYPVDKVSVCVSDDGASMLTFDALAETSEFARKWVP 420
Db 361 TVDPMKEPPIVNTANTVLSILAVDYPVDKVSVCVSDDGASMLTFDALAETSEFARKWVP 420
Qy 421 KKYDIEPRAPYFQKIDYLDKDKVQPSFVKDORRAMKREYEFKIRINALVSKALKVPBE 480
Db 421 KKYNTIEPRAPYFQKIDYLDKDKVHPSFVKDORRAMKREYEFKIRINGLVAKAQVPBE 480
Qy 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLYVYSREKRPQFQHHKAG 540
Db 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLYVYSREKRPQFQHHKAG 540
Qy 541 ANNALVRVSAVLNTNGQYMLNDCDHYINNSKAVREAMCFMDPNLGPQVCYVQFPORFDG 600
Db 541 ANNALVRVSAVLNTNGQYMLNDCDHYINNSKALREAMCFMDPNLGRSVCIYVQFPORFDG 600
Qy 601 IDRDYANRNTVFFDINRLGLDGIQGPVYVGTGCVFNRTALYGYEPPITKAKKPG-FLAS 659
Db 601 IDRDYANRNTVFFDINRLGLDGIQGPVYVGTGCVFNRTALYGYEPPITKQKKGSFLSS 660
Qy 660 LCGGKKKASKRSDDKKSNKHVDSVPVFNLEIDIEGVGAGDFDDEKSLVMSOMSL 719
Db 661 LCGGKKKASKRSDDKKSNKHVDSVPVFNLEIDIEGVGAGDFDDEKSLVMSOMSL 720
Qy 720 KRFGQSAAFVASTLMEYGGVPOSSTPESLLKEAHVISCYEDKSEWGTGEMWIGYSVTE 779
Db 721 KRFGQSAAFVASTLMEYGGVPOSATPESLLKEAHVISCYEDKTEWGTGEMWIGYSVTE 780
Qy 780 DILTGFKMEHARGMRSYVCMKRPAPFKGSAPIMLSRLNQVLRWALGSVEILFSRHCPWY 839
Db 781 DILTGFKMEHARGMRSYVCMKRPAPFKGSAPIMLSRLNQVLRWALGSVEILFSRHCPWY 840
Qy 840 GYGGRLKFLERFAYINTTLYPLTSLPLLYCYILPATCLLTGKFMPEISNLASINFIALF 899
Db 841 GYGGRLKFLERFAYINTTLYPLTSLPLLYCYILPATCLLTGKFIPEISNFAISINFISLF 900
Qy 900 LSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFQGLLKVLAGIDTNTFTVTSK 959
Db 901 LSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFQGLLKVLAGIDTNTFTVTSK 960
Qy 960 ANDEGDPAELMYFKWTTLLIPPTTLLIINMGVAVAGTSYAINSGVQSGGPIFGKLPFAF 1019
Db 961 ASDEGDPAELMYFKWTTLLIPPTTLLIINMGVAVAGTSYAINSGVQSGGPIFGKLPFAF 1020
Qy 1020 WVIVHLYPPLKGLMGRQNTPTTIVVAVILLASIFSLLMVRVDPFTTTRLAGNFIQPCGLN 1079
Db 1021 WVIVHLYPPLKGLMGRQNTPTTIVVAVILLASIFSLLMVRVDPFTTTRVGTQTCGLN 1080
Qy 1080 C 1080
Db 1081 C 1081
```

RESULT 2
Q9LLI1

```
ID Q9LLI1 PRELIMINARY; PRT; 1079 AA.
AC Q9LLI1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cellulose synthase-9.
GN CESA-9.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
ON NCBI_TaxID=4577;
RX MEDLINE=123:1313-1324(2000).
RA "A comparative analysis of the plant cellulose synthase (CesA) gene
RT family."
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL; AF200533; AAF89969.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1079 AA; 120762 MW; EB334272C2ED2DD7 CRC64;

Query Match 93.9%; Score 5423.5; DB 10; Length 1079;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 996; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

Qy 1 MGDADALSKRGHAGDVCCIACDGLGTLTLDGVFTACDVCPRPCPYEHERKEGTOA 60
Db 1 MEGDADGVKSGRRGGQVCQICGDGVGTGTTAGDVTACDVCPRPCPYEHERKDGTGA 60
Qy 61 CLQCKTKYKRHKGSFAIRGEEDGTDADDGSDFNYPASGTEDQKQKIADRMRSWRMNTG 120
Db 61 CPQCKTKYKRHKGSFAIRGEEDGTDADDASDNYPASGNDQKQKIADRMRSWRMNAAG 120
Qy 121 SGNVCHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db 121 SGNVCHPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMSPTGNIGRR 180
Qy 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTIAPSEGRAATI 240
Db 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTIAPSEGRVGD 240
Qy 241 DASTYNNMEDALLNDETQPLSRKVPFIASSKINPYRMVIVLRLVLSIFLHYRLINPVN 300
Db 241 DASTYNNMEDALLNDETQPLSRKVPFIASSKINPYRMVIVLRLVLSIFLHYRLINPVN 300
Qy 301 AYLPLMILLSVICIWFALSILANDYVPDKVSCVSDDGASMLTFDALAETSEFARKWVPF 360
Db 301 AYLPLMILLSVICIWFALSILANDYVPDKVSCVSDDGASMLTFDALAETSEFARKWVP 360
Qy 361 TVDPLKEPPIVNTANTVLSILAVDYPVDKVSVCVSDDGASMLTFDALAETSEFARKWVP 420
Db 361 TVDPMKEPPIVNTANTVLSILAVDYPVDKVSVCVSDDGASMLTFDALAETSEFARKWVP 420
Qy 421 KKYDIEPRAPYFQKIDYLDKDKVQPSFVKDORRAMKREYEFKIRINALVSKALKVPBE 480
Db 421 KKYNTIEPRAPYFQKIDYLDKDKVHPSFVKDORRAMKREYEFKIRINGLVAKAQVPBE 480
Qy 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLYVYSREKRPQFQHHKAG 540
Db 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLYVYSREKRPQFQHHKAG 540
Qy 541 ANNALVRVSAVLNTNGQYMLNDCDHYINNSKAVREAMCFMDPNLGPQVCYVQFPORFDG 600
Db 541 ANNALVRVSAVLNTNGQYMLNDCDHYINNSKALREAMCFMDPNLGRSVCIYVQFPORFDG 600
```

```
QY 601 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 660
DB 601 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 660
QY 661 CGGKKKASKSKSSDKKKSNKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 720
DB 661 CGGKKKASKSKK-GSDKKKSKKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 719
QY 721 RFGQSAFVASTLMEYGGVQPSSTPESLLKEAIIHVISCGYEDKSEWTEIGWIYGSVTD 780
DB 720 RFGQSAFVASTLMEYGGVQPSSTPESLLKEAIIHVISCGYEDKSEWTEIGWIYGSVTD 779
QY 781 ILTGFKMHARGWSVYCMKPRPAKSGAPINLSDRLNOVLRLWALGSVEILLFSRHCPLWYG 840
DB 780 ILTGFKMHARGWSVYCMKPRPAKSGAPINLSDRLNOVLRLWALGSVEILLFSRHCPLWYG 839
QY 841 YGRLKFLERFAVNTTIYPLTSLPLIYVYCIILPAICLLTGKFIIMPEISNLASLWFIALFL 900
DB 840 YGRLKFLERFAVNTTIYPLTSLPLIYVYCIILPAICLLTGKFIIMPEISNLASLWFIALFL 899
QY 901 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLFAVFGGLKVLGIDTNTFTVTSKA 960
DB 900 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLFAVFGGLKVLGIDTNTFTVTSKA 959
QY 961 NDEGDPFAELYMFKWTLLIPPTTILINMVGUVVAGTSYAINSGYQSWGPLFGKLFEAFW 1020
DB 960 SDEGDPFAELYMFKWTLLIPPTTILINMVGUVVAGTSYAINSGYQSWGPLFGKLFEAFW 1019
QY 1021 VIVHLVPLFKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGPNTQTCGINC 1080
DB 1020 VIVHLVPLFKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGPNTQTCGINC 1079

RESULT 3
Q9LLI6 PRELIMINARY; PRT; 1077 AA.
AC Q9LLI6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cellulose synthase-4.
GN CESA-4.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoconostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
  family."
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200528; AAF89964.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; P:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf-ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1077 AA; 120501 MW; 5F960E4AA753E2D6 CRC64;

Query Match 93.8%; Score 5421.5; DB 10; Length 1077;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MDGDADALKSGRHGAGDVQCICADGLGTTLDGVDFTACDVCRFPVCRPCYEHKKEGTQA 60
DB 1 MEGDADGVKSGRRGGVCQICGGVGVTAGDVFAACDVCGFPVCRPCYEHKKEGTQA 60
```

```
QY 61 CLOCKKYKXHHRGSPAIRGEGDDTDADGSDENYPASGTEDEKOKIADRMRSWRMTGG 120
DB 61 CPOCKTKYKXHHKGSPAIRGEGDDTDAD--SDENYLASGNEDEKOKIADRMRSWRMTGG 118
QY 121 SGNVGHKPYGDSGIGLASKYDSGIEIPRGYVPSVNSQMSGIEIPGASPDHMMSPGTGNSRR 180
DB 119 SGDVGRPKYDSGIEGLTKYDSGIEIPRGYIPSVTNSQISGIEIPGASPDHMMSPGTGNIKR 178
QY 181 APPFYVNHSPNPREFSGSIGNVAMKERVDMGKMKQDKGAIPTMTNGTSIAPSEGRAATDI 240
DB 179 APPFYVNHSPNPREFSGSIGNVAMKERVDMGKMKQDKGTIPMTNGTSIAPSEGRGVGDI 238
QY 241 DASTEVNMDALLNDETROPLSRKVPCLASSKINPYRMVIVLRVLVLSIFLHYRLTNPVRN 300
DB 239 DASTDVNMDALLNDETROPLSRKVPCLASSKINPYRMVIVLRVLVLSIFLHYRLTNPVRN 298
QY 301 AYPILWLLSVICETWFAISWILDQFPKWFFINRRETYLDRALRYDREGEPSQLAAVDIFVS 360
DB 299 AYPILWLLSVICETWFAISWILDQFPKWFFINRRETYLDRALRYDREGEPSQLAAVDIFVS 358
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYSDDGASMLTFDALAETSEFARKWVPFV 420
DB 359 TVDPMKEPPIVTANTVLSILAVDYPVDKVSVCYSDDGASMLTFDALAETSEFARKWVPFV 418
QY 421 KKYDIEBRAPEPFYFCOKIDYLDKQVOPSFYKDRAMKREYEEPKIRINALVSKALKYPPE 480
DB 419 KKYNIEBRAPEWFSQKIDYLDKQVOPSFYKDRAMKREYEEPKIRINALVSKALKYPPE 478
QY 481 GWMQDGTWPGNNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVSRKRPFGQHHKKAG 540
DB 479 GWMQDGTWPGNNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVSRKRPFGQHHKKAG 538
QY 541 AMNALVRVSAVLNQGVMNLDCDHYINNSKAVREAMCFLMDPNLGPQVCYVQPPQRPDG 600
DB 539 AMNALVRVSAVLNQGVMNLDCDHYINNSKALREAMCFLMDPNLGRSVYVQPPQRPDG 598
QY 601 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 660
DB 599 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 658
QY 661 CGGKKKASKSKSSDKKKSNKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 720
DB 659 CGGKKKASKSKK-GSDKKKSKKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 717
QY 721 RFGQSAFVASTLMEYGGVQPSSTPESLLKEAIIHVISCGYEDKSEWTEIGWIYGSVTD 780
DB 718 RFGQSAFVASTLMEYGGVQPSSTPESLLKEAIIHVISCGYEDKSEWTEIGWIYGSVTD 777
QY 781 ILTGFKMHARGWSVYCMKPRPAKSGAPINLSDRLNOVLRLWALGSVEILLFSRHCPLWYG 840
DB 778 ILTGFKMHARGWSVYCMKPRPAKSGAPINLSDRLNOVLRLWALGSVEILLFSRHCPLWYG 837
QY 841 YGRLKFLERFAVNTTIYPLTSLPLIYVYCIILPAICLLTGKFIIMPEISNLASLWFIALFL 900
DB 838 YGRLKFLERFAVNTTIYPLTSLPLIYVYCIILPAICLLTGKFIIMPEISNLASLWFIALFL 897
QY 901 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLFAVFGGLKVLGIDTNTFTVTSKA 960
DB 898 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLFAVFGGLKVLGIDTNTFTVTSKA 957
QY 961 NDEGDPFAELYMFKWTLLIPPTTILINMVGUVVAGTSYAINSGYQSWGPLFGKLFEAFW 1020
DB 958 SDEGDPFAELYMFKWTLLIPPTTILINMVGUVVAGTSYAINSGYQSWGPLFGKLFEAFW 1017
QY 1021 VIVHLVPLFKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGPNTQTCGINC 1080
DB 1018 VIVHLVPLFKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGPNTQTCGINC 1077

RESULT 4
Q84M43 PRELIMINARY; PRT; 1073 AA.
ID Q84M43
```


AC Q84M43;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative cellulose synthase catalytic subunit.
GN OSUNBA0059E14.5.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanek S., Riedmuller S.B., Uteback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA0059E14 genomic sequence.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135958; RAP21426.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose synth.
DR Pfam; PF03552; Cellulose synt. 1.
SQ SEQUENCE 1073 AA; 120681 MW; 55B677033E77A7BE CRC64;

Query Match 89.5%; Score 5169.5; DB 10; Length 1073;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 950; Conservative 67; Mismatches 54; Indels 5; Gaps 4;

QY 6 DALKSGRGAGDVCOICADGLGTLTLDGVTACDVRCPVCYEHKKEGTQACLOCK 65
DB 2 DKAQSGKQ -CHVCQCGDGVTAADGELFTACDVGFPVCPYERKKGDSQACPQCK 59
QY 66 TKYKRRHSGPALRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTGSGNGV 125
DB 60 TKYKRRHSGPPILGDESDDVADSDVNYPTSGNQDHKHKIAERMLTWKNSGRDDIV 119
QY 126 HPKYSGEIGLSKYDSGEIPRGYPVSVNSQMSGEIPGASPDHMSPTGNISRRRA -PFP 184
DB 120 HSKYDSGEIGHPKYDSGEIPRTIYIPSLTHSQISGEIPGASPD -HMSVPVGNIGRRGHPPP 178
QY 185 VYNHSPNPSRFPSSGSGNVAWKERVDMKODKGAIPMTNCTSTAPSEGRAATDIDAST 244
DB 179 VYNHSPNPSRFPSSGSGNVAWKERVDMKODKGAIPMWANGTSAPSEGRGVGDIDAST 237
QY 245 EYNMEDALLNDETROPLSRKVP IASSKINPYRMVILRLVLSIFLHRLTNFVRNAYPL 304
DB 238 DYNMEDALLNDETROPLSRKVP IASSKINPYRMVILRLVLSIFLHRLTNFVRNAYPL 297
QY 305 WLLSVICSIWALSILWDPKFWPFINRETYLDRALRYDREGSPQAAVDIFVSTVDP 364
DB 298 WLLSVICSIWALSILWDPKFWPFINRETYLDRALRYDREGSPQAAVDIFVSTVDP 357
QY 365 LKEPPIVANTVLSILAVDYPVDKYSVYSDDGASMLTFDALAESEFARKWVPVKYD 424
DB 358 MKEPPIVANTVLSILAVDYPVDKYSVYSDDGASMLTFDALAESEFARKWVPVKYD 417
QY 425 IEPRAPEFYCQKIDYLDKQVPSEVKDORRAMKREYEFKIRINALVSKALVPBEGWIM 484
DB 418 IEPRAPEFYCQKIDYLDKQVPSEVKDORRAMKREYEFKIRINALVSKALVPBEGWIM 477
QY 485 ODGTWPNGNTRDHPFMIQVFLHSGGLDTEGNELPRLVYVYREKRGFGHKKAGAMNA 544

DB 478 ODGTWPNGNTRDHPFMIQVFLHSGGLDTEGNELPRLVYVYREKRGFGHKKAGAMNA 537
QY 545 LVRVSAVLITNGOYMLNLDCDDHYNNKSKAVREACFLMDPNLGPVYCYVQFPORFGIDRN 604
DB 538 LVRVSAVLITNGOYMLNLDCDDHYNNKSKALREACFLMDPNLGPVYCYVQFPORFGIDRN 597
QY 605 DRYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPPIKAKKPGFLASLCCGK 664
DB 598 DRYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPPIKAKKPGFESSLCCGR 657
QY 665 KXASKSKRSDKKKSHKHVDSSVPFNLEDEIEBGEVAGFDDKSVLMSQMSLKRFGQ 724
DB 658 KXASKSKRSDKKKSHKHVDSSVPFNLEDEIEBGEVAGFDDKSVLMSQMSLKRFGQ 717
QY 725 SAAFAVASTLMEYGVQSSPTESLKLKHAHVLSGCCYEDKSEWGTGIMTYGVSVDILTGT 784
DB 718 SSVFVASTLMEYGVQSSPTESLKLKHAHVLSGCCYEDKSDWGTGIMTYGVSVDILTGT 777
QY 785 FPMHARGMRSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCHCPLWYGYGR 844
DB 778 FPMHARGMRSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCHCPLWYGYGR 837
QY 845 LKFLERPAYINTIYPLTSLPLLYCILPAICLLTGKIMPRIINSLASIWFTALFLSIPA 904
DB 838 LKFLERPAYINTIYPLTSLPLLYCILPAICLLTGKIMPRIINSLASIWFTALFLSIPA 897
QY 905 TGILEMRWSGVGIDSEWRNEQFWVIGGISAHLFAVFGQLLKVLAGIDTNFTVTSKANDEE 964
DB 898 TGILEMRWSGVGIDSEWRNEQFWVIGGISAHLFAVFGQLLKVLAGIDTNFTVTSKANDEE 957
QY 965 GDFASLYMFKWTTLLIPPTTILINMGVWAGTSYAINSGYQSGWGLFGKLFPAFWVIVH 1024
DB 958 GDFASLYMFKWTTLLIPPTTILINMGVWAGTSYAINSGYQSGWGLFGKLFPAFWVIVH 1017
QY 1025 LYPFLKGLMGRQRTPTTIVWVALLASIFSLWLVRVDPFTRLAGPNIQTGGINC 1080
DB 1018 LYPFLKGLMGRQRTPTTIVWVALLASIFSLWLVRVDPFTRTRVGPDTQKCGINC 1073

RESULT 5
Q9LLI5 PRELIMINARY; PRT; 1076 AA.
AC Q9LLI5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cellulose synthase-5.
GN CESA-5.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoconostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene family.";
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200529; AAF89965.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose synth.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF03552; Cellulose synt. 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1076 AA; 120724 MW; 41EE7C20EDA54F27 CRC64;

Query Match 89.0%; Score 5143; DB 10; Length 1076;
Best Local Similarity 87.6%; Pred. No. 0;

Matches	947;	Conservative	69;	Mismatches	59;	Indels	6;	Gaps	6
QY	1	MDGDADALKSRGHGAGDYCOICADGLTGLTDGDVFTACDVCRFPPVCRPCYBERHEKGTQA	60						
Db	1	MDG-GDATNSGKHVAGVCQICGDGVGTAADGLFTACDVCGFVCRPCYBERKDGTQA	59						
QY	61	CLOCKTKYKRHSIPAIRGEEDDTADGSDFNYPASGTEDQKQIADRMVSRWMNTGG	120						
Db	60	CPQCKTKYKRHKGSPPHVGENEODVADVDYNYQASGNQDQKQIAERMLTWRTNSRG	119						
QY	121	SGNVGHPKYDSEIGLSKYDSEIIPRGVYPSVWNSQMSGEIPGASPHHMSPTGNTLSRR	180						
Db	120	S-DIGLAKYDSEIGHGKYDSEIIPRGVIPSUTHISQISEIPGASPD-HMSPVGNIGRR	177						
QY	181	A-PPFVNVHSPNPSREFSGSIGNVAKWERVDGWKMKDGAIPTMNTGTSIAPSGRAATD	239						
Db	178	GHQFPVNVHSPNPSREFSGSLGNVAWKERVUDGKWK-DKGAIPTMNTGTSIAPSGRGVAD	236						
QY	240	IDASTEYNMEDALLNDETROPLSRKVPITASSKINPYRMVIVLRUVLVISFLHYRLTNPVR	299						
Db	237	IDASTDYNMEDALLNDETROPLSRKVPITPSSSRINPYRMVIVLRVAVLVCIFLRYRITHPNV	296						
QY	300	NAYPLWLLSVICETWEALSWILOQFPKWPNNRRTYLDRLALRYDRGERPSQLAAVDIFV	359						
Db	297	NAYPLWLLSVICETWEALSWILOQFPKWPSPNNRRTYLDRLALRYDRREGTSQAPVDIFV	356						
QY	360	STVDPLKEPPIVTANTVLSILLADVPYVDKVCYVSDDGASMLTFDALAETSEFARKWVPF	419						
Db	357	STVDPMKEPPLVTANTVLSILLADVPYVDKVCYVSDDGAAWLPFDALSETSEFARKWVPF	416						
QY	420	VKKYDIEPRAPEFFCQIDYLDKQVQSPFVKDORRAMKREYEPKIRINALVSKALKVPE	479						
Db	417	CKYNIIEFPAPEWYFAQIDYLDKQVQTSFVKERRAMKREYEPFKVRINGLVAKAQKVP	476						
QY	480	EGWTMQDQTPWGNNTRDHPGMIQVFIHGSGGLDTEGNELPRIVYVSREKRPQFQHHKKA	539						
Db	477	EGWTMQDQTPWGNNTRDHPGMIQVFIHGSGGLDVEGNELPRIVYVSREKRPQFQHHKKA	536						
QY	540	GAMNALVRVSAVLNTGOYMLNLDCDHYINNSKAVREAMCFLMDPNLPQVCYVQFQRPD	599						
Db	537	GAMNALVRVSAVLNTGOYMLNLDCDHYINNSKALREAMCFLMDPNLGRNVICYVQFQRPD	596						
QY	600	GIDENDRYANRNTVFFDINIRGLDGIQGPVYVGTGCYVFNRTATLYGYEPPKAKKQPGFLAS	659						
Db	597	GIDENDRYANRNTVFFDINIRGLDGIQGPVYVGTGCYVFNRTATLYGYEPPVKKKCPGFSS	656						
QY	660	LCGGKKKASKSKRSKDKSNKHVDSSVPVENLDEIEEGVEGAGNDEKSVLMSQMSLE	719						
Db	657	LCGGRKKTYSKK-SSEKKSHRHADSVPVFNLEDIEEGIQSGQDDEKSLIMSQMSLE	715						
QY	720	KRFQSSAAFAVSTIMEYGGVPQSSPESLLKEAHIVISCGYEDKSBWGTEIGHIYGSVTE	779						
Db	716	KRFQSSVFFVASTIMEYGGVPQSATPESLLKEAHIVISCGYEDKTWGTTEIGHIYGSVTE	775						
QY	780	DILITGFQKHARGWRSVTCMPKRPAPKGSAPINISDLRNQVLRWALGSVEILFSRHCPWY	839						
Db	776	DILITGFQKHARGWRSIYTCMPKRPAPKGSAPINISDLRNQVLRWALGSIELFSRHCPWY	835						
QY	840	GYGRULKFLERFAVINTTIYPLTSLPLLYVCIIPALCLLTGKIMPEINLASIWETIALP	899						
Db	836	GYGRULKFLERFAVINTTIYPLTSLPLLYVCIIPAVCLLTGKFIIPKISLNLSVMPISLF	895						
QY	900	LSIFATGILEMRWSGVGIDEMWNEQFPWTGGISAHLFAVQGLLKVLAGIDNFTVTSK	959						
Db	896	ISIFATGILEMRWSGVGIDEMWNEQFPWTGGISAHLFAVQGLLKVLAGIDTFTVTTSK	955						
QY	960	ANDEEGDFAELYNFKWTTLLIPPTTLIIINMVGVAGTSYAINSGYQSGWGLFGKLFFAF	1019						
Db	956	ATDEEGDFAELYNFKWTTLLIPPTTLIIINLVGVVAGISVAINSGYQSGWGLFGKLFFAF	1015						
QY	1020	WVIVHLVPLFKGLMGRONRTPTIIVWVALLASIFSLLVWRVDVDPFTRLAGNIOQTCGIN	1079						
Db	1016	WVIVHLVPLFKGLMGRONRTPTIIVWVALLASIFSLMVWRIDPFTTRVTDIAKCGIN	1075						

QY		1080 C 1080			
DB					
DB		1076 C 1076			
RESULT 6					
Q9XGK6 PRELIMINARY; PRT; 1067 AA.					
ID	Q9XGK6;				
AC	Q9XGK6;				
DT	01-NOV-1999 (TrEMBLrel. 12, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE	Cellulose synthase catalytic subunit.				
GN	CELA3.				
OS	Gossypium hirsutum (Upland cotton).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.				
OX	NCBI_TaxID=3635;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RP	STRAIN=cv. Texas marker-1; TISSUE=Fiber;				
RX	MEDLINE=20098737; PubMed=10631273;				
RA	Laosinchai W., Cui X., Brown R.M. Jr.;				
RA	"A full length cDNA of cotton cellulose synthase has high homology				
RT	with the Arabidopsis RSM1 gene and the cotton Celal gene (Accession				
RT	No. AF200453) (PGR 00-002).";				
RL	Plant Physiol. 122:291-291(2000).				
RN	[2]				
RC	SEQUENCE FROM N.A.				
RP	STRAIN=cv. Texas marker-1; TISSUE=Fiber;				
RA	Kimura S.; Laosinchai W.; Itoh T.; Cui X.; Brown R.M. Jr.;				
RT	"Immunogold Labeling of Rosette Terminal Cellulose Synthesizing				
RT	Complexes in a Vascular Plant (Vigna angularis).";				
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF150630; AAD39534.2; -				
DR	GO; GO:0016020; C.membrane; IEA.				
DR	GO; GO:0016760; F.cellulose synthase (UDP-forming) activity; IEA.				
DR	GO; GO:0030244; P.cellulose biosynthesis; IEA.				
DR	InterPro; IPRO00150; Cellulose_synth.				
DR	InterPro; IPRO01841; Znf ring.				
DR	Pfam; PF03552; Cellulose_synt; 1.				
DR	SMART; SM00184; RING_1_-synth.				
SQ	SEQUENCE 1067 AA; 119325 MW; 0BBA2ED00590F29C CRC64;				
Query Match 79.2%; Score 4577.5; DB 10; Length 1067;					
Best Local Similarity 78.6%; Pred. No. 0;					
Matches 853; Conservative 100; Mismatches 105; Indels 27; Gaps 11					
QY	2	DGDADALSKGRHGAGDVCOIICADGLGTLLDGVFTACDVCRFPCVCPVEHKEGTQAC 61			
DB	4	EGDIGG-KPMKNLGGQTQICGDNVNKNTDGDFFTACNICAFPCVPCPYEERKDGNQSC 62			
QY	62	LQCKTKYKRHRGSPAIRGEEDTDADDG-SDFNVPSGTEDOKIADRMSWRMTGG 120			
DB	63	PQCKTRYKWQSGFALLGDRETGGDADDGASDFY--SENQEKQKLARMQGWNAKYR 120			
QY	121	SGNVGHPKYDSGETGLSKYDSGETPRGYVPVTSNQ-MSGEPGASPDHHMSPGTGNISR 179			
DB	121	GEDVGAPTYDK-----EISHNHPIPLLTSGQEVSGELSAAAPERLSMASPGVAGG 169			
QY	180	RAPPEYNHSPNSREF-SGSIGNVANKERVDGWKKODKCAIPMTNGTSTAPSEGRAAT 238			
DB	170	KSSIRVV----DPVREPGSGLGNAVANKERV DGWKKQEKNTVPMs--TCOATISE-RGLG 222			
QY	239	DIDASTEYNMEDALLINDETROPISRKYVIPASSKINPVRMWIVLRVLVISIFLHVRLTNPV 298			
DB	223	DIDASTDLVDQSQNDEARQLSRKYVSSSKINPVRMWIILRLVILCIFLHVRLTNPV 282			
QY	299	RNAYPLMLLSVICHIWFALSWILDQFPKPFPINRETLDRLALRYDREGESQLAADVF 358			
DB	283	RNAYPLMLLSVICHIWFALSWILDQFPKLPVNRETLDRLALRYDREGESQLAADVF 342			

[illegible][illegible]

Matches	733; Conservative	138; Mismatches	168; Indels	39; Gaps	15;
Qy	2	DGDADALXSGRHGAGDVQCI	CADGLGTLDGVTACDVCRFPVCRPCYEHKEGTQAC	61	
Db	25	DGGTKPLKMN--GQICQICDDVGLA	TGTVFVACNECAFVCRPCYEHKEGTQCC	81	
Qy	62	LOCKTKYKRRHSGPAIRGEGDDTD	DGSDNFNPASGTEDQKQIADMRMRWMTGGS	121	
Db	82	POCKTRFRHRHSPRVGEGDEDD	VDDIENEFNYAGCANKARHQHGEFFS-----	133	
Qy	122	GNVGHPKYDSGEIGLSKYD---	SGEIPRGYVPSVNSQMSGEBIPGASPDHMHMSPTGNIS	178	
Db	134	---SSRHESQPIPLTHGHTVS	GEIRTPDTQSVRTT--SGEL-GPSDRNAISSP--YTD	184	
Qy	179	RRAPPYVNHNSPNREP--SGSIG	NVANKERVDSGKQKQKAI	PWNTNGTSTAPSEGRAA	237
Db	185	PROQPEV--RIYDPSKDLNSY	GLGNVDKERVGKMLQKQKMLQWTG----	KYHEGKG	238
Qy	238	TDIDASTEYNMEDALLNDET	ROPGRKVPFIASKINPYRMVILRLVLSIFLHYRLTMP	297	
Db	239	-BIE-GTSGNGELQWAD	TRLPMSEVVPISRLTPYRVVILRLILLCFLQYRTTHP	296	
Qy	298	VRNAVPLWLSVICETW	FALSWMILDQPKWFPINRETYDLRALRYDRGEBFSQLAVDI	357	
Db	297	VKNAYPLWLSVICETW	FALSMILDQPKWFPINRETYDLRALRYDRGEBFSQLPVVDV	356	
Qy	358	FVSTVDPLKEPPIV	TANTVLSILAVDYPVDKYSVYDDGASMLT	PDALAEETSEPAKWV	417
Db	357	FVSTVDPLKEPPIV	TANTVLSILAVDYPVDKYSVYDDGASMLT	PDALAEETSEPAKWV	416
Qy	418	PVWKYDI	EPRAPEFFCQIKIDYLDKQVPSFKDRRANKREYEEFKIRINALVSKALKV	477	
Db	417	PFCKENI	EPRAPEFFCQIKIDYLDKQVPSFKERRANKREYEEFKIRINALVAKAQKI	476	
Qy	478	PEEGTMDQGT	PMGNTRDHPQMIQVFLHSGGLDTEGNE	LPRLVYVREKRPQHHK	537
Db	477	PEEGTMDQGT	PMGNTRDHPQMIQVFLHSGGLDTEGNE	LPRLVYVREKRPQHHK	536
Qy	538	KAGAMNALRVSAVL	TNGQYMLNLCDDHYINNSKAVREAMCF	FLMDPNLGPVCYVFPOR	597
Db	537	KAGAMNALRVSAVL	TNGAYLLNVDCDHPNNSKAIKEAMCF	MDPAIGKCCYVFPOR	596
Qy	598	FDGIDRNDRYANNT	VFFDINLRGLDGIQGPVVVGTGCVFNRTAIYGYPPPIKAK--KPG	655	
Db	597	FDGIDLHDYANRNI	VFFDINLMKGLDGIQGPVVVGTGCCFNQALYGYDPLTEEDLEPN	656	
Qy	656	FLA-SLCGGKKKASK	KKKSSDKKXKNHVDSSVPVFNLEIDEGVEGAGFDDKSVLMS	714	
Db	657	IIVKSCGSRKKGSK	KKYNYEKRRGINRSDSNAPLFNMEDIDEGFE--GYDDERSILMS	714	
Qy	715	QMSLEKREFGOSAF	VASTLMGYGVPOSSTPESILKEAIIHVISCGYEDKSEWTEIGWIY	774	
Db	715	QRSVEKRFQOSPF	IAATFMEQGGIPPTINPATLLKEAIIHVISCGYEDKTEWKEIGWIY	774	
Qy	775	GSVTEIDILTFKMHARG	WRSVYCMPPKPAFKGSAPINLSRLNQVLRWALGSEVILLFSRH	834	
Db	775	GSVTEIDILTFKMHARG	WISYIYCNPPPAFKGSAPINLSRLNQVLRWALGSEVILLSRH	834	
Qy	835	CPWYGYGGRKLE	RFAPYINTIYPLTSLPLLYCYILPAICLLTKFIMPEISNLASIW	894	
Db	835	CPWYGYGGRRLER	IAINTIYPIITSPLIAYCILPAFLITDRFIPEISNYSIW	894	
Qy	895	FIALFISIFATGILEMR	WSGVGIDEMWRNEQFWVIGGISAHLEFAVFGILLKVLGADITNF	954	
Db	895	FILLFISIAVTGILELR	WSGVSDWWRNEQFWVIGGTSAHLEFAVFGILLKVLGADITNF	954	
Qy	955	TVTISKANDERDGA	EALYMFKWTLLIPPTTILINMVGVVAGTSYAINSGSQSGWPLFGK	1014	
Db	955	TVTISKATDEGD	FAELYIFKWTALLIPPTTVLLVNLIGIVAGSYAVNSGYSQSGWPLFGK	1014	
Qy	1015	LFFAFVWIVLHYPFL	KLMGRONTPTIVWAVILLASIFSLLMVRYDVPFTTRLAGPN	1072	
Db	1015	LFFALWIAHLHYPFL	KGLGRONTPTIVVWSVILLASIFSLLMVRIINPFVD--ANPN	1070	

RESULT 12

Q9LLI2

ID Q9LLI2

AC Q9LLI2;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Cellulose synthase-8.

CESA-8.

Zea mays (Maize).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

NCBI_TaxID=4577;

[1]

SEQUENCE FROM N.A.

RX MEDLINE=20398328; PubMed=10938350;

RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,

RA Xoonostle-Cazares B., Delmer D.P.;

RT "A comparative analysis of the plant cellulose synthase (Cesa) gene

family";

RT Plant Physiol. 123:1313-1324(2000).

RL EMBL; AF200532; AAF89968.1; --

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.

DR GO; GO:0030244; P:cellulose biosynthesis; IEA.

DR InterPro; IPR005150; Cellulose_synth.

DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF03552; Cellulose_synt; 1.

DR SMART; SM00184; RING; 1.

DR PROSITE; PS00089; ZF_RING 2; 1.

SQ SEQUENCE 1094 AA; 122575 MW; DFBBC18F49D23F5E CRC64;

Query Match

Best Local Similarity

Matches 723; Conservative

138; Mismatches 168; Indels 39; Gaps 15;

Score 3844.5; DB 10; Length 1094;

Pred. No. 1.3e-301;

Indels 69; Gaps 15;

Qy 9 KSGRHGAGDV-----CQICADGLGTTLDGVTACDVCRFPVCRPCYEHKEGTQACL 62

Db 25 ESGAGGGAARAEAPCQICGDEGVFGDEFPVACNECAFVCRACIYERREGSQACP 84

Qy 63 QCKTKYKRRHSGPAIRGEGDDTDADGSDFN--YPASGTEDQKQIADMRMRWMTGGS 121

Db 85 QCRTRYKRLKGCPRVAGDEEDGVDLEGEFGLQDGAHEDDPQYVAESMLRAQMSYGRG 144

Qy 122 GNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVNSQMSGEIPGASPDHMHMSPT-----G 175

Db 145 GD-AHP-----GFSPPVNPVPLTNGQWDDIP---PEQHALVPYSMSGGG 186

Qy 176 NISRRAPPFPVYN-----HSPNPSREFSG--SIGNVANKERVDSGKMKQDKGAIPMTNGT 227

Db 187 GCKRIHPLFADPNLPVQPSRMDPSKDAAAYGYSVANKERMESGWKQKQER-----L 238

Qy 228 STAPSSEGRAATDIDASTEYNMEDALLNDETROPQLSRKVPFIASSKINPYRMVILRLVLS 287

Db 239 QHVRSEGGGDWDCDDA-----DLPLMDEARQPLSRKVPISSSRINPYRMVILRLVVLG 292

Qy 288 IFLHYRLTWPVRNAYPLWLSVICETWALSILAVDYPVDKVSICYSDGASMLTFDALL 407

Db 293 FFHYRVMHAPAKDAFWLWLSVICETWAFNSWIDQFPKWLPIERETYLDRLSRFPDKEG 352

Qy 348 EPSQLAAVDIFVSTVDPLKEPPIVTANTVLSILAVDYPVDKVSICYSDGASMLTFDALL 407

Db 353 QPSQLAPIDFFVSTVDPTKEPPLVTANTVLSILSVDPVEKVSICYSDGAAAMLTFPALS 412

Qy 408 ETSEFARKWVPFKKYDIEBRAPEFFPCQIKIDYLDKQVQSPFYKQDRAMKREYEEFKIRI 467

Db 413 ETSEFAKKWVPFKFNIEBRAPEFWFQKIDYLDKQVQSPFYKQDRAMKREYEEFKIRI 472

Qy 468 NALVSKALKVPESGWIWMDGTPWPGNTRDHPGMQVFLHSGGLDTEGNEPLRVYSR 527


```
Db 473 NALVAKAQVPEEGWTMDGSPWPGNNVRDHPGMIQVFLGSGGRDVEGNELPLRVVYSR 532
Qy 528 EKRFQFQHHKAGAMNALRVSAVLITNGQYMLNLDCHYINNSKAVREAMCFMDPNLGP 587
Db 533 EKRFQYHHKAGAMNALRVSAVLNSAAYLLNLDCHYINNSKAIKEAMCFMMDPLVGGK 592
Qy 588 QVCYVQFPQRFDDGDRNDRYANRNVTFDDINLRGLDGIQGPVYVGTGCVFNRTAIYGEY 647
Db 593 KVCYVQFPQRFDDGDKNDRYANRNVTFDDINMKGLDGIQGPVYVGTGCVFRQALYGYDA 652
Qy 648 PIKAKKP-----GFLASLCGKKGKASKSKRSDDKKKS--NKHVDSVVPFVNL 695
Db 653 P-KTKKPPSRTCNCPKWCWCLSCCCKSRNKKKTKTKPKTEKKKLFKKAENPSPAYALGE 711
Qy 696 IEBCVEGAGDEDEKSVLMSQMSLEKRFQOSAAFAVASTIMEYGGVQPSSTPESILKEAIV 755
Db 712 IDEGAPGA--DIEKAGIVNQOKLEKFGQSSVFVASTLLENGGTLKSASPASLKEAIV 769
Qy 756 ISCYEDKSWGTETIGYISVTEIDITLGFQKHARGWRSVYCMKPRPAFKGSAPINLSDR 815
Db 770 ISCYEDKTDWKEIGMIYGSITDITLGFQKHCHGWSYICIPKRPAPKGSAPINLSDR 829
Qy 816 LNOVLWALGSVEILFRSHCPLWYGYGRKLFERFAYINTIYPLTSLPLVYCTLPAL 875
Db 830 LHQVLRWALGSVEIFFSKHCHFLWYGYGGKLFERFSYINSIVYVWTSIPLLAYCTLPAL 889
Qy 876 CLLTGKTEIMPEISNLASIWFIATLFSATGILEMRWSGVGIDEWNRNEQFWIGTSAH 935
Db 890 CLLTGKFTPELTNVAISWFIATLFSATGILEMRWSGVGIDEWNRNEQFWIGTSAH 949
Qy 936 LFVAFQGLLVKLAGIDNTFTVTSKANDBEQDFEALYMFKWTLLIPLPTTILINMGVVA 995
Db 950 LFVAFQGLLVKAVAGIDTSFTVTSKAGDDE-EFSELYTFKWTLLIPLPTTILINMGVVA 1008
Qy 996 GTSVAINSGYQSWGFLPGKLFPAFWIVHLYPFLKGLMGRQNRPTTIVIVAVALLSIFS 1055
Db 1009 GISNAINNGYESWGFLPGKLFPAFWIVHLYPFLKGLVGRQNRPTTIVIVMSILLASIFS 1068
Qy 1056 LLWVRVDPFTTRLAGPNIQCGINC 1080
Db 1069 LLWVRVDPFLAKSGPLLEBEGDUC 1093

RESULT 13
Q851L8
ID Q851L8 PRELIMINARY; PRT; 1092 AA.
AC Q851L8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cellulose synthase.
GN CSUNE0042109.26.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teitron T., Kim M.M., Bera J.J., Jin S.S.,
RA Radrosh D.W., Tallon L.J., Koo H., Zismann V., Heisao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RA "Oryza sativa chromosome 3 BAC OSJNbA0042109 genomic sequence.";
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR ENMBL; AC044487; AA041140.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
```

```
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 1092 AA; 123395 MW; A682D0AA434BE24D0 CRC64;

Query Match 66.3%; Score 3833; DB 10; Length 1092;
Best Local Similarity 65.1%; Pred. No. 1.1e-300;
Matches 719; Conservative 147; Mismatches 177; Indels 62; Gaps 17;

Qy 2 DGDADALKSRHSGADVQCICADGLGTLTLDGVFTACDVCRFPVPCRCYERHERKEGTQAC 61
Db 23 DGEFGP-KPVKHTNGQVCQICGDDVGLTPDGEPPVACNECAFFVCRDCYEREREGTQNC 81
Qy 62 LOCKTKYKRRGSPAIRGEGDDTDADD-GSDFNYPASGTEDQKQKADMRWRMNTGG 120
Db 82 PQCKTRFKRLKGCARVPGDE-EEDVDLLENFNM--RDKTDQYVAESMLHGMYSYGR 137
Qy 121 SGNV-GHPKYDSGIEGLSKYDSGIEPRGYVPSVTNSOMSGEIPGASPDHMMST--GN 176
Db 138 GGDLDGVQHFQ-----PIPNVPLLTNGEMADDIP--PEQHALVPSFMGG 181
Qy 177 ISSRAPFPYVN-----HSPNPSREFSG-SIGNVAWKERVDMKQDKGAIPTMGTS 228
Db 182 GKRIHPLPYADENLPVQPSMDPSKDLAAYGYGSAVKERMESWKQKQERLHQMRNDG-- 239
Qy 229 IAPSEGRAATDIDASTEYNMEDALLDETQPLSRKVPDIASSKINPYRMVTLVLVLSI 288
Db 240 -----GGKMDGDDGDD----ADLPLMDEARQPLSRKIPISSSLVNPYRMIILRVVLGF 290
Qy 289 FLHYRLNPNVRNAVPIMLLSVICEIFALSWILDQFPKWPINRETYLDRALARYDREGE 348
Db 291 FFHYRVWHPVDFAPALMLISVICEIFAMSWILDQFPKWPFIETTYLDRUTLRFDEKG 350
Qy 349 PSQALAAVDIFVSTVDPKPEPIVTANTVLSILAVDYPDKVSCYVSDGASMLTFDALAE 408
Db 351 QSQALAVDFVSTVDPKPEPLVTANTVLSILAVDYPDKVSCYVSDGASMLTFEALSE 410
Qy 409 TSFPAKWWVFKKYDIEPRAPFYYFOKIDYLDKQVQPSFVKDRRAMKREYEFKIRIN 468
Db 411 TSFPAKWWVFKKYSLEPRAPFYYFOKIDYLDKQVAPNPFVRRERRAMKREYEFKIRIN 470
Qy 469 ALYSKALKVPEEGWIMODGTPWPGNNTRDHPGMIQVFLGSGGLDTEGNEPLRVVYSRE 528
Db 471 ALVAKAQVPEEGWTMDGSPWPGNNVRDHPGMIQVFLGSGGHVDEGNELPLRVVYSRE 530
Qy 529 KRPFQFQHHKAGAMNALRVSAVLITNGQYMLNLDCHYINNSKAVREAMCFMDPNLGP 588
Db 531 KRPFQYHHKAGAMNALRVSAVLITNAPYMLNLDCHYINNSKAIKEAMCFMMDPLVGGK 590
Qy 589 QVCYVQFPQRFDDGDRNDRYANRNVTFDDINLRGLDGIQGPVYVGTGCVFNRTAIYGEY 648
Db 591 QVCYVQFPQRFDDGDRNDRYANRNVTFDDINMKGLDGIQGPVYVGTGCVFRQALYGYDAP 650
Qy 649 IKAKKPGFLASLC-----GGKKKASKSKRSDDKKKS--NKHVDSVVPFVNL 695
Db 651 -KSKKPPSRTCNCPKWCWCLSCCCKSRNKKKTKTKPKTEKKKLFKKAENQSPAYALGE 709
Qy 696 IEBCVEGAGDEDEKSVLMSQMSLEKRFQOSAAFAVASTIMEYGGVQPSSTPESILKEAIV 755
Db 710 IDEGAPGA--ENEKAGIVNQOKLEKFGQSSVFVASTLLENGGTLKSASPASLKEAIV 767
Qy 756 ISCYEDKSWGTETIGYISVTEIDITLGFQKHARGWRSVYCMKPRPAFKGSAPINLSDR 815
Db 768 ISCYEDKTDWKEIGMIYGSITDITLGFQKHCHGWSYICIPKRAAFKGSAPINLSDR 827
Qy 816 LNOVLWALGSVEILFRSHCPLWYGYGRKLFERFAYINTIYPLTSLPLVYCTLPAL 875
Db 828 LHQVLRWALGSVEIFFSNHCPMLWYGYGGKLFERFSYINSIVYVWTSIPLLAYCTLPAL 887
Qy 876 CLLTGKFTPEISNLASIWFIATLFSATGILEMRWSGVGIDEWNRNEQFWIGTSAH 935
Db 888 CLLTGKFTPELTNVAISWFIATLFSATGILEMRWSGVGIDEWNRNEQFWIGTSAH 947
```

```
QY 936 LFAVQGLLKVLAGIDTNTFTVTSKANDEBGFADLYMEKWTLLIPPTTILIIINMGVVA 995
Db 948 LFAVQGLLKVLAGIDTFTVTSKGGDE-EFSELYTFKWTLLIPPTTILIIINMGVVA 1006
QY 996 GTSVAINGYSQSWGLPGLKLFPAFVWVHLVPLKGLMGQRNRTTIVIVWAVLLASIFS 1055
Db 1007 GVSNAINGYESWGLPGLKLFPAFVWVHLVPLKGLVGRNRTTIVIVWSILLASIFS 1066
QY 1056 LLWVRVDPTTFLAGPNIQTGGINC 1080
Db 1067 LLWVRIDPFLAKNDGPLEEBCGLDC 1091

RESULT 14
Q9LLI4 PRELIMINARY; PRT; 1059 AA.
AC Q9LLI4;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
D3 Cellulose synthase-6.
GN CESA-6.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xocoostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
RT family";
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200530; AAF89966.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; P:cellulose biosynthesis; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1059 AA; 120044 MW; 289DA26B2532249 CRC64;

Query Match 66.1%; Score 3820; DB 10; Length 1059;
Best Local Similarity 65.1%; Pred. No. 1.2e-299;
Matches 714; Conservative 140; Mismatches 168; Indels 74; Gaps 16;

QY 16 GDVCOICADGLGTLDGVDFTACDVCPPVPCYEHKEGTQACLOCKTKYKGRGSP 75
Db 6 GQVCGIGDDVGRNDFGPFVACNECAFICRDCYEYERREGTQCPQCKTRFKFKGCA 65
QY 76 AIRBEGDDTDADGSDFNYPASGTEDQKQIADRMRSMNMTGGSNVG-----HPKY 129
Db 66 RVPGBDEEDGVLDLEFNFW----SKHDSQYLAESMLHAHMSYGRGADLDGVPPQHP-- 120
QY 130 DSGEIGLSKYDGETPRGVVPVTSVMSGSEIFGASPDHMMMSPT---GNISRRAPFPYV 186
Db 121 -----IPN--VPLTNGQVDDIP---PQHALVPFSGVGGGRIIHLPLPYA 161
QY 187 N-----HSPNPSREFSG-STGNVAKERVQDKMKQDKGAIPMTNGTSIAPSEGRAAT 238
Db 162 DNLVPQPSMPSDKDLAAYGSGVAKERMESWQKQER--MHQTR-----NDGGDD 213
QY 239 DIDASTEYMEDALLNDETQPLSRKVP IASKINPYRMVIVLRLVLSIFLHYRTNPV 298
Db 214 GDDA-----DLPLMEARQPLSRKIPLPSSQINPYRMIIIRLVLCFFHYRVMHPV 266
QY 299 RNAPYPLWLSVCEIWFALSILDQPKWFPINRETYLDRLALRYDREGSPQLAAVDIF 358
Db 267 PDAFALWLSVCEIWFAMSWILDQPKWFPFERETYLDRLSLRFDKEGHPQLAPVDFF 326
```

```
QY 359 VSTDPLKEPPIVTANTVLSILAVDPVDKVS CVYSDDGASMLTFDALAETSEFARKWVP 418
Db 327 VSTDPLKEPPIVTANTVLSILSVDPVDKVS CVYSDDGAAMLTFEALSETSEFARKWVP 386
QY 419 FVKKYDIDPRAPEFVFCQKIDYLDKQVOPSVFKDRAMKREYEEFKIRINALVSALKVP 478
Db 387 FCKRYSLEPRAPEFWYFQKIDYLDKQVAPNFVRERRAMKREYEEFKIRINALVAKAQVP 446
QY 479 BEGTMQDGTWPNGNTRDHPGMIQVFLGSHSGGLDTEGNEPRLVYVREKRPFGPHKK 538
Db 447 BEGTMQDGTWPNGNVRDHPGMIQVFLGQSGHDVEGNEPRLVYVREKRPFGNHHKK 506
QY 539 AGAMNALVRVSAVLNQTGYMLNLDCHYINNS KAVREAMCFMLDPNLGPQVCYVFPQRF 598
Db 507 AGAMNALVRVSAVLNAPYLLNLDCHYINNS KAIKEAMCFMMDPLLGGKVCYVFPQRF 566
QY 599 DGDIDRYANRNTVFFDINRLGLDGIQGVVVGTCVFNRTAIYGYBPPIKAKP---- 654
Db 567 DGDIDRYANRNVVFFDINMKLDGIQGVVVGTCVFRQALYGYDAP-KTKKPPGRT 625
QY 655 -----GFLASLGGKKKASKSKRSSDK-----KSKNHVDSVPVFNLEDIEEGVEGAG 704
Db 626 CNCWFKWICCCCFGNRTKTKTKSKPKFKIKL FKKNQAPAYALGEIDEAAPGA- 684
QY 705 FDDEKSVLMSQMSLEKRFQSAAFVASTIMEYGGVPOQSTPESLLKEAIIHVISCYEDKS 764
Db 685 -ENEXASIVNQKLEKFKQSFSVFVASTLLENGGTLKGSASPASILLKEAIIHVISCYEDKT 743
QY 765 EWGTETGHWYGSVTEIDITGFKMHARGHRSVYVCMKRPFAFKGSAFINSLDRNLNOVLWAL 824
Db 744 GWGKDIGHYGSVTEIDITGFKMHCHGWRYSYICIPKRAAFKGSAPLNSDRHQVLRWAL 803
QY 825 GSVELFSRKHCPWYGYGGRKFLERFAYINTIYPLTSLPLLVCILPAICLLTGKFTM 884
Db 804 GSIEILFSNHCPWYGYGGLKFLERFSYINSIVVPWTSIPLLAYCTLPALICLLTGKFT 863
QY 885 PEISNLASIFWLFALFSLPATGILEMRWSGVGIDEWNRNEQFWITGGISAHUFAVPGLL 944
Db 864 PELNNVASLWFMSLFICIFATPSILEMRWSGVGIDWNRNEQFWITGGVSSHLLFAVPGLL 923
QY 945 KVLAGIDNFTVTSKANDEBGFADLYMEKWTLLIPPTTILIIINMGVWAGTSVAINS 1004
Db 924 KVIAGVDSFTVTSKGGDE-EFSELYTFKWTLLIPPTTILIIINMGVWAGTSVAINS 982
QY 1005 YOSWGLPGLKLFPAFVWVHLVPLKGLMGQRNRTTIVIVWAVLLASIFSLLWVRVDPF 1064
Db 983 YESWGLPGLKLFPAFVWVHLVPLKGLVGRNRTTIVIVWSILLASIFSLLWVRIDPF 1042
QY 1065 TTRLAGPNIQTGGINC 1080
Db 1043 LAKDDGPLEEBCGLDC 1058

RESULT 15
Q9LLI3 PRELIMINARY; PRT; 1086 AA.
AC Q9LLI3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cellulose synthase-7.
GN CESA-7.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xocoostle-Cazares B., Delmer D.P.;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 17:17:01 ; Search time 6177 Seconds
(without alignments)
5221.168 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVQC.....VDPFTRLAGNIQTGINC 1080

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPto_spool/US0900037/runat_18082004_081514_14648/app.query.fasta_1.1223
-DB=EST -QFMT=fastcap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0900037@cgn_1_4365@runat_18082004_081514_14648 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=10 -XGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_esti:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5421.5	93.8	3763	11	AY108113	Zea mays
2	4937.5	85.5	3728	11	AY112236	Zea mays
3	4805	83.2	3696	11	AY110079	Zea mays
4	3902	67.5	3898	11	AY110415	Zea mays
5	3817	66.1	3783	11	AY104730	Zea mays
6	3806	65.9	3897	11	AY103655	Zea mays
7	3805	65.9	3788	11	AY103701	Zea mays
8	3525	61.0	2872	11	AY104236	Zea mays
9	2507	43.4	1874	14	CD726831	EST027_Cu
10	1575	27.3	916	14	CK272603	EST18681
11	1542	26.7	924	14	CF513822	CABud0007
12	1530.5	26.5	977	14	CK271294	EST17372
13	1503.5	26.0	958	14	CK274934	EST721012
14	1463.5	25.3	978	12	BG321258	Zm04_03f1
15	1430	24.7	953	14	CD439980	EL01N0531
16	1429.5	24.6	874	14	CD433620	EL01N0313
17	1422	24.6	865	14	CD440398	EL01N0554
18	1418.5	24.6	1189	11	AY107656	Zea mays
19	1407	24.4	880	12	BM816138	HCI09812
20	1390.5	24.1	925	14	CF213555	CGF100079
21	1376	23.8	875	14	CF514581	CABud0005
22	1359	23.5	1014	29	CG203358	PUCB90TB
23	1356.5	23.2	966	14	CA247044	SCCCAM209
24	1341	23.0	913	14	BG368813	AF53-Rpf
25	1321.5	22.9	870	12	BG368813	HVSME1002
26	1314.5	22.8	1022	28	BZ827769	PUPH017TB
27	1313	22.7	807	12	B1311236	EST531298
28	1310.5	22.7	960	9	AB049602	AB049602
29	1307	22.6	762	13	BQ802778	WHE2829_H
30	1306	22.6	781	14	CF447738	EST684083
31	1294.5	22.4	837	14	CF512600	CABud0003
32	1290	22.3	809	14	CF445367	EST681712
33	1289.5	22.3	905	13	CA159025	SCZR2312
34	1289	22.3	840	13	CA151353	SCFZR2200
35	1286	22.3	810	14	CF444535	EST680880
36	1282	22.2	750	10	BF624748	HVSME001
37	1274	22.0	1042	14	CA239344	SCBFF1507
38	1273	22.0	851	14	CB892280	EST643249
39	1271	22.0	1133	14	CK208799	FGAS02052
40	1268	21.9	807	14	CA918393	EST642540
41	1264	21.9	821	29	CG383072	OG1BB95TV
42	1261	21.8	741	10	BF460078	072G01_Ma
43	1259	21.8	827	14	CF437366	EST673711
44	1256	21.7	701	14	CD930401	GR45_111C
45	1251.5	21.7	816	14	CB097089	AF53-Rpf_

ALIGNMENTS

RESULT 1
AY108113
LOCUS AY108113
DEFINITION Zea mays PC0126465 mRNA sequence.
ACCESSION AY108113
VERSION AY108113.1 GI:21211191
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3763)

AUTHORS Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,
TITLE Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
JOURNAL Maize Mapping Project/DuPont Consensus Sequences for
REFERENCE Overgo Probes
AUTHORS 2 (bases 1 to 3763)
TITLE Coe, E.H.
JOURNAL Direct Submission
COMMENT Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSL, maizenap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.

FEATURES
 source
 1. 3763
 /organism="Zea mays"
 /mol_type="rRNA"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed DuPont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3763
 Score: 5421.50 Matches: 998
 Percent Similarity: 96.67% Conservative: 46
 Best Local Similarity: 92.41% Mismatches: 33
 Query Match: 93.83% Indels: 3
 DB: 11 Gaps: 2

US-09-900-237A-30 (1-1080) x AY108113 (1-3763)

QY	1	MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyValGlyAspValCysGln	20
DB	338	ATGGAGGGGACGCGGACGGGTGAAGTCCGGGAGCGCGGTGGCGGACGAGTGTGCCAG	397
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
DB	398	ATCTCGCGGACGGCGTGGCCACCACGCGGAGGGGACGTCTTCGCGCGCTCGGACGTC	457
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGlyThrGlnAla	60
DB	458	TGCGGGTTTCGGGTGTGCGCGCCCTGCTACGAGTACGAGCGCAAGACGCGCAGGCG	517
QY	61	CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
DB	518	TGCCCCCATGTGACAGACCAAGTACAGCCGACCAAGGGGAGCCCGCGCATCCGTGGGGAG	577
QY	81	GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
DB	578	GAAGGAGACACACTCATGCCGAT-----AGCGACTTCAATTAACCTTGCATCTGGCAAT	631
QY	101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly	120
DB	632	GAGGACCAAGCAGAGAGATGCGGACAGAAATCGGCAGTCTGGCGCATGACGTGGGGGC	691
QY	121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
DB	692	AGCGGGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGAC	751
QY	141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160

Db	752	AGTGGCGAGATTCTCCGGGATACATCCCATCAGTCACACAGCCAGATCTCAGAGAA	811
QY	161	IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg	180
Db	812	ATCCCTGTGTCTCCCTGACCATCATATGATGCCCACTGGGAAATATGGCAGCGT	871
QY	181	AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle	200
Db	872	GCTCCATTCCCTATGTGAACCATTCGCCAAATCCGTCAAGGAGTCTCTCTGTAGCAT	931
QY	201	GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla	220
Db	932	GGGAATGTTCCCTGGAAGAGAGAGGTTCATGCTGCGAAATGAACGACGACAAAGGCG	991
QY	221	IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle	240
Db	992	ATTCCCATGACGAATGGCACAAGCAATGCTCCTCTGAGGGTCGGGGTCTTGGTGATAT	1051
QY	241	AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro	260
Db	1052	GATGCATCAACTGATTACAACATGGAAGATGCTTATTGAACGACGAAACTCGACAGCT	1111
QY	261	LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal	280
Db	1112	CTATCTAGGAAGTTCCACTTCCTCCAGGATAATCCATCAGGATGGTCATGTG	1171
QY	281	LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn	300
Db	1172	CTGGCATTTGTTCTAAGCATCTCTGCACTACCTGATCACAATCTCTGGCGCAT	1231
QY	301	AlaTyrProLeuTrpLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle	320
Db	1232	GCATACCCATTATGGCTTCTATCTGTATATGTGAGATCTGTTGCTCTTTCGTGGATA	1291
QY	321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
Db	1292	TTGGATCAGTTCCTCAAGTGGTTTCCAATCAACCGGAGACGACGTCTGATAGCTGGCA	1351
QY	341	LeuArgTyrAspArgGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer	360
Db	1352	TTAAGTATGACGGGAAGTGAGCATCTCAGTTGGCTGCTGTGACATTTTCGTGAGT	1411
QY	361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380
Db	1412	ACAGTCGACCCCAATGAAGGAGCTCTCTCTGTCTCACTGCCAATACCGTGTATCCATTCT	1471
QY	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1472	GCTGTGATTTACCTGTGATGAAGTCTCTTGTATGTATCTGTATGATGATGAGCTCGATG	1531
QY	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
Db	1532	CTGACATTTGATGACATAGCTAGACTTCAGACTTTCAGAAATGGGTACCATTTGTT	1591
QY	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1592	AAGAAGTACAACATTCGAAGCTCTGATGCTGATGCTCTCTCCAGAAAAATGATTAC	1651
QY	441	LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr	460
Db	1652	TTGAAGGACAAAGTGCACCTTCATTTGTTTAAAGACCGCGGCGCATGAGAGAGATAT	1711
QY	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1712	GAAGATTCAAAGTTAGGTTAAATGGCTTTGTTGCTAAGGCACAGAAAGTTCTCTGAGAA	1771
QY	481	GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	500
Db	1772	GGATGGATCATGCAAGATGGCACCATGCGCCAGGAAACAAATACCMGGGACCATCTCGGA	1831
QY	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1832	ATGATTTCAGGTTTCTCTGTGTCACAGTGTGGCTTGATCTAGGGGCATGAGTACCC	1891

Qy	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly	540	Qy	881	LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu	900
Db	1892	CGTTGGTCTATGTTCTCGTGAAGAAGCTCTGAGATCCAGCATCACAGAAGAACTGCT	1951	Db	2969	AAGTTCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC	3028
Qy	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560	Qy	901	SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr	920
Db	1952	GCCATGAATGCTCTCTGTTCTCGTGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAT	2011	Db	3029	TCGATCTTCGCCACGGGCATCTCGAGATGAGTGGAGGGGGTGGGATCGACAGATGG	3088
Qy	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580	Qy	921	TyrArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe	940
Db	2012	CTTGATTTGTGATCACTACATTAACACAGATGAGGCTCTCAGGGAAGCTATGTGCTTCCTT	2071	Db	3089	TGGAGAAACGAGCATCTTCGGTGTATCGGGGGCATCTCCGGCACCTCTTCGCCGTGTC	3148
Qy	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600	Qy	941	GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla	960
Db	2072	ATGGACCTTAACCTAGGAAGAGTGTCTGCTACGTCCAGTTCCTCCAGAGATTCGATGGC	2131	Db	3149	CAGGCGCTTGCTCAAGAGTGTGGCGGGCATCGACCAACTTCACCGTCACCTCCCAAGGCC	3208
Qy	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620	Qy	961	AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle	980
Db	2132	ATTGACAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTCGATATTAACTTGAGA	2191	Db	3209	TCGGACGAGGACGGGCATCTCGCGGAGCTGTACATGTTCAAGTGGACGACGCTCTTGATC	3268
Qy	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640	Qy	981	ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla	1000
Db	2192	GGTCTTGATGATCCCAAGGACCGATTTATGTGGAACCTGCTGTGTTTCAACCGAACA	2251	Db	3269	CGGCCACACCATCTGATCATCAACTGGTCGCGTCTCGCGGCATCTCTTACGCC	3328
Qy	641	AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu	660	Qy	1001	IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyr	1020
Db	2252	GCTCTATATGTTATGACCCCAATTAAGCAGAAGAAGGTGTTCTTGTCAATCACTA	2311	Db	3329	ATCAACAGCGATACCAAGTCTGTGGGGCCGCTCTTCGCAAGCTCTTCTCGCTTCTGG	3388
Qy	661	CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer	680	Qy	1021	ValIleValHisLeuTyrPyrPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro	1040
Db	2312	TGTGCGGTAGGAAGAAGCAAGCAAAATCAAAGAAG---GGCTCGGACAAGAAGATCG	2368	Db	3389	GTCATCGTCCACCTGTACCCGTTCTCAAGGGCTCTATGGGCGAGGACGACACCCCG	3448
Qy	681	AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal	700	Qy	1041	ThrIleValIleValTyrAlaValLeuLeuAlaSerIlePheSerLeuTyrValArg	1060
Db	2369	CAGAAGCATGTGGACAGTCTGTGCGCAGTATTCACCTTGAAGATATAGAGAGGAGTT	2428	Db	3449	ACCATCGTGTGCTGTGGCCATCTCTGTGGCGTCCATCTTCTCTGTGTGGTTCCG	3508
Qy	701	GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys	720	Qy	1061	ValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys	1080
Db	2429	GAAGCGCTGTGATTGACGACGAGAATCACTTTATGTCTCAATGAGCTGTGGAGAG	2488	Db	3509	ATCGACCCCTTCCACCCCGCTCACTGGCCCGGATATCCAGACGTGTGGCATCAACTGC	3568
Qy	721	ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro	740	RESULT 2			
Db	2489	AGATTGTGCCAGTCCGACGCTTTGTTGCCCTCCACTCTGATGGAGTATGGTGTCTCT	2548	LOCUS	AV112236	3728 bp	mRNA linear HTC 17-OCT-2002
Qy	741	GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr	760	DEFINITION	Zea mays CL1160_1	mRNA sequence.	
Db	2549	CAGTCCCGCACTCCGGAGTCTCTTCTGAAGAAGCTATCCATGTTATAAGCTGTGGCTAT	2608	ACCESSION	AV112236		
Qy	761	GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp	780	VERSION	AV112236.1	GI:21216826	
Db	2609	GAGGACAAGACTGAATGGGNACTGAGATCGGTGGATCTACGTTCTGTGACAGAAGAC	2668	KEYWORDS	HTC.		
Qy	781	IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys	800	SOURCE	Zea mays		
Db	2669	ATTCTACCGGATTCAGATGACCGCGAGGCTGGCGTCTGATCTACTGATGCCCAAG	2728	ORGANISM	Zea mays		
Qy	801	ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu	820	REFERENCE	Unpublished (2002)		
Db	2729	CGGCGAGCTTCAAGGGGTGCGCCCAATCAATCTTTCGGAACCGTCTGAACAGGTGCTC	2788	REFERENCE	2 (bases 1 to 3728)		
Qy	821	ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly	840	TITLE	Direct Submission		
Db	2789	CGTGGGCTCTTGGGTCTGGTGAAGATCTCTTCAGCGCGCACTGCCCTGTGTGTACGGC	2848	JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of		
Qy	841	TyrGlyGlyArgLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	860	COMMENT	Missouri, Columbia, MO 65211, USA		
Db	2849	TACGAGGGCGGCTCAAGTTCCTGGAGATTCGCGTACATCAACACCATCTACCCG	2908		If you are interested in getting corresponding physical clones,		
Qy	861	LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly	880		these are publicly available from ZmDB and may be found by BLAST		
Db	2909	CTCAGTCCATCCGCTTCTCATCTACTGATCTGCGCGCCATCTGTCTGCTCACCAGA	2968		searching at MSI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,		
					www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the		
					maize cDNA sequences is either Virginia Walbot, Stanford or Pat		
					Schnable, Iowa State, then clones may be requested from ZmDB:		
					www.zmdb.iastate.edu.		
					Location/Qualifiers		
					1. -3728		
					/organism="Zea mays"		


```

/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0 Length: 3728
Score: 4937.50 Matches: 918
Percent Similarity: 89.07% Conservative: 44
Best Local Similarity: 85.00% Mismatches: 117
Query Match: 85.45% Indels: 1
DB: 11 Gaps: 1

US-09-900-237A-30 (1-1080) x AY112236 (1-3728)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
238 ATGGAGGGGCGACGGCGCGCGTGAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGCCAG 297

QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 ATCTCGCGGCGATGGCGTGGCGCACTACGGCGGAGGAGAGCGTCTTCACCGCCTGCGACGTC 357

QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlyThrGlnAla 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
358 TCGCGGTTCCCGGTGGCGCCCTCTACGAGTACGAGCGCAAGACGGCACACAGCG 417

QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 TGNNNNNAGTGCNNNNNCAAGTACAGCGCCACAGCGGAGTCCAGCGATCCGAGGGGAG 477

QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 GAAGGAGACGATACTGATCGCATGATGCTAGCGACTTCAACTACCTGCTGCTGCGCAAT 537

QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
538 GACGACCAAGAGAGAGAGATGTGTGACAGATCGCGAGTGGCGCATGATGCTNNNNNC 597

QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
598 AGCGGGGATGTTGGCGCGCCCAAGTATGACAGTGTGTGAGATCGGGCTTACCAAGTACGAC 657

QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
658 AGTGGTGAGATCCCTCGGGGATACATCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTCGGGAGAA 717

QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
718 ATCCCTGGGTGCTTCCCTGACCATATATATGATGCTCTTACTCGGGAACATTCGACGGCGC 777

QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778 GCCCCATTTCCCTATATGAATCATTCATAAATCCGTCGAGGGAATTCCTCTGATGCGGTT 837

QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
838 GGGAAATGTTCCCTCGGAAAGAGAGGTTGATGGCTGGAAAATGAAGCAGGACAAAGGGAACA 897

QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
898 ATTCCCATGCAAGATGGCAACAGCATTTGCTCCCTCTGAGGGGCCGGGTGTGTTGATATT 957

QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

DB 958 GATGCATCAACTGATTACAAACATGAAGATGCTTTATTAAACGATGAACACTGCCAGCCT 1017
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1018 CTATCTAGAAAGTTCACCTTCTCTCCAGATAAATCCATACAGATGTCATTGTG 1077

QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1078 CTACATTGATTGTTCTAAGCATCTTCTTGCACTACCGGATCACAATCTCTGGGTAAT 1137

QY 301 AlaTyrProLeuTyrPheLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle 320
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1138 GCATACCCCACTGTGGCTTCTATCTGTATGTGAGATCTGGTCTGCTCTTCCCTGGATA 1197

QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1198 TTGGATCAGTTTCCAAAGTGGTTTCCAAATCAACCGCGAGACTTACTTGTATAGATCGCA 1257

QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1258 TTAAGGTATGACCGGGAAGGTGAGCCATCTCAAGTTGGCTGTGTTGACANNNNNGTCAGT 1317

QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1318 ACTGTGACCCCAATGAAGAGCTCTCTGTGCACATGCGCAATACCGTGTCTATCCATCTC 1377

QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 GCTGTGGACTATCTCTGTGGATAAGGTCTCTGCTATGATCTGTATGATGATGAGCTGCTATG 1437

QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1438 CTGACATTTGATGCTAGCTAGACTTCAGAGATTGTGTAGAAAAATGGTGGCTTGTGT 1497

QY 421 LysLysTyrAspIleGluProGluAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1498 AAGAAGTACAACTTGAACCTAGAGCTCTGTAATGGTACTTCTCCAGAAAAATGATTATAC 1557

QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaAlaMetLysArgGluTyr 460
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1558 TTGAAGGACAAAGTGCACCTTCATTTGTTAAAGCCGCGGNNNNNNNNNNNNNNNNNN 1617

QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1618 NNNNNNNCAAATAATTAGGCTAAATGGCTTGTGTGTAGGCAACNNNNNGTCCCTGAGGAA 1677

QY 481 GlyTrpIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1678 GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATACACGAGGACCAATCTGGA 1737

QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1738 ATGATTTCAGTTTCTCTGTCACAGTGTGTGCTTGTATCTGATGAGGGTAAATGAGCTACCC 1797

QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1798 CGTTTGTCTATGTTCTCTGNNNNNGTCTCTGATTCAGCATCACAGAAAGTGGT 1857

QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1858 GCCATGAATGCTCTGTGCGGCTCTCAGTGTGCTTACCAATGGAACAAATACATGTTGAAT 1917

QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1918 CTTGATTGTGATCATTACATCAACACAGTAAAGCTCTCAGGGAAGCTATGTGCTTCTT 1977

QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1978 ATGCATCTTAACTAGGAAGAGTGTCTGCTATGTTTCAAGTTTCCCGCAGAGTTCGATGGT 2037

QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2038 ATTGATAGGAATGATCGATATGCCAACAGGAACACCGTGNNNNCGATATTAACTGAGA 2097

```

```
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 2098 GGTCTTGTATGGCATCAAGACGACAGTTTATGTGGGCACCTGGCTGTGTTTCAACAGAAC 2157
QY 641 AlalIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2158 GCTCTATATGGTATATGAGNNNNAAATTAAGCAAAAGAGGGTGGTTTCTTGTGCATCACTA 2217
QY 661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db 2218 TGTGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2274
QY 681 AsnLysHisValAspSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2275 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2334
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluTyr 720
Db 2335 GAAAGCGCTGGATTTGATGATGAGAAATCACTTCTTATGTCTCAATGAGCTTGAGAG 2394
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2395 AGATTGGCCAACTGCGAGCTTTGTGGTCCACTCTGATGGAATNNNNNNNNNNNNNNNN 2454
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2455 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2514
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrGlySerValThrGluAsp 780
Db 2515 GAGGACAAGATTGATGGGAACTGAGATTGGTGGATCTATGTTCTGTGACGGAAGAT 2574
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2575 ATTCTACGTGGTTCAAGATGCACGACGAGCTGGCGTGCATCTACTGCTAGTGGCTAAG 2634
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2635 CGGCGGCGCTTCAAGGATCGCTCCCATCAATCTCTCAGACCGTCTGACAGGTGCTC 2694
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrGly 840
Db 2695 CGGTGGGCTCTCGGTTCAGTGGAAATCCTTTTCAGCGGCGCATGNNNNNTATGGTACGG 2754
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2755 TACGAGGACCGCTGAAGTTCTTGAGAGATTGCGCTACATCAACACACCATCTACCCG 2814
QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2815 CTCACGTCCCTCCGCTCCTCATTTACTGTATCTGCTGCATCTGCTGCTCAGCGG 2874
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2875 AAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGTATCTGGTTCACTCTCTCTTCATC 2934
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 2935 TCGATCTTCGCCACCGGTATCTCGAGATGAGGTGGAGCGCGCTGGCGCATCGACGAGTGG 2994
QY 921 TrrArgAsnGluGlnPheTrrValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2995 TGGAGGAACGAGCATGTTCTGGTTCATCGAGGCATCTCCGCCACCTCTTCGCCGTCTTC 3054
QY 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3055 CAGGGCTCTCTCAAGTGGCTTGGCGGCATCGACACCAACTTCACCGTCACCTCCAGGCC 3114
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrrThrThrLeuLeuIle 980
Db 3115 TCGGATGAAGACGCGGACTTTCGCGGAGCTGTACATGTTCAAGTGGAGCAGACATCTCTGATC
```

```
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3175 CGCCACACCATCTGATCATCAACCTGGTGGCGTGTGTTCGCGGCATCTCTCTACGCC 3234
QY 1001 IleAsnSerGlyTyrGlnSerTrrGlyProLeuPheGlyLysLeuPhePheAlaPheTrr 1020
Db 3235 ATCAACAGCGGTACACAGTCTGGTGGGTTCGCTCTTCGGCAAGCTCTTCCTCGCTTCGG 3294
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3295 GTGATCGTTTCACTGTACCCGTTCCTCAAGGGTCTCATGGTCCGCGAGACCGCACCCG 3354
QY 1041 ThrIleValIleValTrrAlaValLeuLeuAlaSerIlePheSerLeuLeuTrrValArg 1060
Db 3355 ACCATCGTGGTGTCTGGCGCATCTGCTGGCGTGCATCTCTCTGCTGTGGTTCGC 3414
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3415 ATCGATCGTTTCAACCAACCGGTCTACTGCCCGGATACTCGAACGTGTGGCATCACTGC 3474

RESULT 3
AY110079 3696 bp mRNA linear HTC 17-OCT-2002
LOCUS Zea mays CL1164_1 mRNA sequence.
DEFINITION Zea mays
ACCESSION AY110079
VERSION AY110079.1 GI:21214162
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3696)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 3696)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
location/Qualifiers
1..3696
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630055"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3696
Score: 4805.00 Matches: 895
Percent Similarity: 89.08% Conservative: 68
Best Local Similarity: 82.79% Mismatches: 112
Query Match: 83.16% Indels: 6
DB: 11 Gaps: 6
```

```
US-09-900-237A-30 (1-1080) x AY110079 (1-3696)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 254 ATGGACGGC--GGCGACGCCACGAATTCGGGGAAGCATGTGGCGGCGAGGTGTGCCAG 310
QY 21 IleCysAlaAspGlyLeuThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 311 ATCTCGCGCACGGCGTGGCACCGCGCGGACGGCGACCTCTTCAACGCCCTGCGACGTC 370
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlyThrGlnAla 60
Db 371 TCGCGCTTCCCCGTGTCGCCCATCTGCTACGAGTACGAGCGACGACGGCACCCACGGCG 430
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 431 TCGCCGCGAGTGAAGTAACTAAGCGCCACAAAGGGAGGCCACCAAGTACACGGGTGAG 490
QY 81 GluGlyAspAspThrAspAlaAspGlySerAspPheAsnTyrProHlaSerGlyThr 100
Db 491 GAAATGAGAGATGTGATGCTGACGATGTGAGTGAATCACTCACTCACTCACTCACTCACT 550
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db 551 CAGGATCAGACGAAAGATTCGTGAGAGAACTGCTCCTTGGCGGACAACTCACTGCGC 610
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 611 AGT---GATATTGGCCTGGCTAAGTATGACAGCGGTGAAATTGGCGATGGGAAGTATGAC 667
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 668 AGTGGTGAATCCTCGTGATATATCCCGTCACTAACTCACTCACTCACTCACTCACTCACT 727
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 728 ATTCCTGGAGCTTCCCTCGAT---CATATGATGTCTCTGTTGGGAACATTCGACGGCT 784
QY 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 785 GGACATCAATTCCTTATGTAAATCATCTCCAAACCCATCGAGGGAGTTCTCCGGTAGC 844
QY 200 IleGlyAsnValAlaIlePlysGluArgValAspGlyTyrLysMetLysGlnAspLysGly 219
Db 845 CTTGGCAATGTTCATCGAAAGAGGGTGGATGGATGGAAATGAAG---GATAAAGGT 901
QY 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
Db 902 GCATTCCTATGACCAATGGACACAGCATTTGCTCCATCAGAAGGGCGTGGAGTTGCTGAT 961
QY 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 962 ATTGATGCTTCTACTGATTATAACATGGAAGATGCCTTACTGAATGATGAATCTCGCAA 1021
QY 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
Db 1022 CCTCTATCTAGAAAAGGCCAATTCCTTCATPCAGAAATAATCCGTACAGAAATGTCATT 1081
QY 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
Db 1082 GTGCTAGCTTGGCTGTCTATGCATATCTTGGCTACCGTACCATCACTCACTCACTCACT 1141
QY 300 AsnAlaTyrProIleThrLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTyr 319
Db 1142 AATGCATATCCACTGTGGCTTTTATCCGTATATGTGAGATCTGGTTTGTCTGTCTG 1201
QY 320 IleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeu 339
Db 1202 ATTTTGGATCAGTTCCCAAGGTGTCCTCCCAATCAACCGTGAACATACCTTGTATGACTG 1261
QY 340 AlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaValAspIlePheVal 359
```

```
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2399 AAGAGATTGGCCAGTCCAGTGNNNNGTAGCCTCTACTCTGATGAATATGGTGTGTT 2458
QY 740 ProGlnSerSerThrProGluSerLeuLeuLeuGluAlaHisValIleSerCysGly 759
Db 2459 CCACATCTGCAACTCCNNNNNNNNNNNNNNNNNNNNNTATTTCATCTCATCAGCTGTGC 2518
QY 760 TyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu 779
Db 2519 TATGAGCAAAACTGACTGGGAACCTGAGATTGGGTGGATCTATGGTCTCTGTACAGAA 2578
QY 780 AspileLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro 799
Db 2579 GACATTCTCACCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCT 2638
QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2639 AAGCAGCAGCAGCTTTCAAGGGATCTGCTCTATCAACCTTCGATCGTTGAATCAAGTG 2698
QY 820 LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr 839
Db 2699 CTTCGGTGGCTCTTGGTTCATGAAATCTTTTCAGCAGCATGTGCCATATGGTAT 2758
QY 840 GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2759 GGCTATGGAGCGCGCTTAAATCTCTGGAGAGATTGCTTATATCAACAACAATTTAT 2818
QY 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2819 CCACCTCATCAATCCGCTCTCTCTGACTGATATGCCAGCATTTGTCTTCTCACT 2878
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPhe 899
Db 2879 GGAAGTTCATCATCCCAAGATTAAGTAACTAGAGAGTGTGTGGTTTATATCGCTCTTT 2938
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyLeuAspGlu 919
Db 2939 ATCTCAATCTTGGCAGTGTATCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNCT 2998
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 2999 TGGTGGAGAAACGACGACTCTGGGTCACTGTGGTATTTCTCGCAATTTATTTGGCGTC 3058
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAnPheThrValThrSerLys 959
Db 3059 TTCAGGCTCTCCTGAAGGTGCTTGTGGTATCGACACGAGCTTCACCTGCACCTCTAAG 3118
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrTrpThrLeuLeu 979
Db 3119 GCCACTGACGAAGAAGGTGATTTTGGCAGCTCTACATGTTCAAGTGGACACGCTTCTG 3178
QY 980 IleProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3179 ATCCCAACCAACCACTATTTTTCATCAACCTCGGCGTGTGCTGGCATTTCCTAC 3238
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3239 GCAATCAATAGCGGTACCAGTCATGGGGACCTCTTTTCGGGAAGCTCTTCTTTGGCTC 3298
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3299 TGGGTGATTGTCCACCTGTACCCCTTCTCAAGGGCCCTCATGGGAGAGCAGAACCGCAG 3358
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3359 CCGACCATTTGTGTGTGTGGGTATCTCTCTCGTCGATCTTTTCCCTGATGTGGGT 3418
QY 1060 ArgValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3419 CGTATCGATTCATTCACACCGGGTCACTGGTCTCTGATATCGCGAAATGTGGCATCAAC 3478
```

```
QY 1080 Cys 1080
Db 3479 TGC 3481
RESULT 4
LOCUS AY110415
DEFINITION Zea mays CULL166_1 mRNA sequence.
ACCESSION AY110415
VERSION AY110415.1 GI:21214824
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3898)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3898)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1..3898
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630057"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

```
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3898
Score: 3902.00 Matches: 737
Percent Similarity: 79.27% Conservative: 131
Best Local Similarity: 67.31% Mismatches: 169
Query Match: 67.53% Indels: 58
DB: 11 Gaps: 22
US-09-900-237A-30 (1-1080) x AY110415 (1-3898)
```

```
QY 2 AspGlyAspAlaAspAla---LeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 255 GACGGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 315 ATTTCGGCGACCTCTGTGGCGGTTTCAGCCCATCTGGTGATCTCTTTGTTGCTGCAATGAG 374
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 375 TGTGCTTCCCTGCTGCGCCCTTGTATGATGATGACGAGCGCAAGGAGGAGGAGCAATGCG 434
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
```



```

QY 789 AlaAqGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla 808
Db 2535 NNNAGGGGCTGCATCAATCACTACGTGATGCACACACACCTTTTCAAGGGTTCTGCA 2594
QY 809 ProIleAsnLeuSerAspArgLeuAsnGlnValLeuAArgTyrAlaLeuGlySerValGlu 828
Db 2595 CCAATCAATCTTCTGATGCTTAAATCAGGTGCTCCGTTGGGCTTTGGGTCACTGGAA 2654
QY 829 IleLeuPheSerArgHisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeu 848
Db 2655 ATTCTGCTTAGCAGACATGCTCTATCTGGTATGTTTAAATGGACGATTAAGCTTTTG 2714
QY 849 GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuVal 868
Db 2715 GAGAGCTGGCTTACATCAACACCATGTTTATCCATCAACATCTGTCGCTTAACGCC 2774
QY 869 TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluLeuSer 888
Db 2775 TATTGTGCTCTCTGCTATCTCTCTTACCAATAAATTTATCATCTCTGAGATTAGC 2834
QY 889 AsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu 908
Db 2835 AATTATGCTGGAGTGTCTTCAATCTCTTTCGCTCCANNNGCCACTGGTATATG 2894
QY 909 GluMetArgTyrSerGlyValGlyIleAspGluTyrTyrArgAsnGluLysPheTyrVal 928
Db 2895 GAGCTTAGATGAGTGTGTTGGCATTAAGATTTGGTGGAGAAATGACAGTTTGGTT 2954
QY 929 IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuValLeuAla 948
Db 2955 ATTGTGTGCACCTCTGCCATCTCTTCGACGTGTTCCAGGTCTGCTGAAAGTGTGCT 3014
QY 949 GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGlyAspPheAla 968
Db 3015 GGGATTGATACCAACTTACAGTTTACCTCAAAGGCATCTGATGAGGATGGGACTTTGCT 3074
QY 969 GluLeuTyrMetPheLysTyrThrThrLeuLeuIleProProThrThrIleLeuIle 988
Db 3075 GAGCTATATGTTCAAGTGACAGATTTGCTCATCTCTCCGACACATGTTCTTGTCATT 3134
QY 989 AsnMetValGlyValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyr 1008
Db 3135 AACCTGTGCGAATGTTGGCAGGAATTTCTGATGATGCAATTAACAGTGGCTACCACTCTCG 3194
QY 1009 GlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPhe 1028
Db 3195 GGTCCGCTCTTTGAAAGCTTCTTCTCGATCTGGGTGNTCTCCATCTGTACACGTTTC 3254
QY 1029 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTyrAlaVal 1048
Db 3255 CTCAGGGTCTCATGGAGAGCAGAACCGCACCAACATTCGTCATTTGCTGTGTCATC 3314
QY 1049 LeuLeuAlaSerIlePheSerLeuLeuTyrValArgValAspProPhe-----Thr 1065
Db 3315 CTTCTTGCACTATCTTCTCTCTGCTGTTGGGTGAAGATCGATCTTTCATCTCCCGACA 3374
QY 1066 ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyLysLeuAsnCys 1080
Db 3375 CAGAAAGCTGCTGCC-----TTGGGGCAATGTGGGCTCACTGC 3413

RESULT 5
AY104730
LOCUS AY104730 3783 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays P0100501 mRNA sequence.
ACCESSION AY104730
VERSION AY104730.1 GI:21207808
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.

```

```

REFERENCE 1 (bases 1 to 3783)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
1..3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3783
Score: 3817.00 Matches: 714
Percent Similarity: 77.92% Conservative: 140
Best Local Similarity: 65.15% Mismatches: 168
Query Match: 66.06% Indels: 74
DB: 11 Gaps: 16
US-09-900-237A-30 (1-1080) x AY104730 (1-3783)
QY 16 GlyAspValCysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPhe 35
Db 291 GGCCAGGTGTGCCAGATTTCGGCGCAGCAGCTGGGGCGCAACCCGACGGGAGCGGTC 350
QY 36 ThrAlaCysAspValCysArgPheProValCysArgProCysTyrGluHisGluArgLys 55
Db 351 GTGGCCTGCACGAGTGGCGCTTCCCATCTGCGGGAGCTGCTACGAGTACGAGCGCGC 410
QY 56 GluGlyThrGlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerPro 75
Db 411 GAGGGCAGCAGAACTGCCCCAGTCGCAAGCCGCTTCAAGCCCTCAAGGGTGGCGC 470
QY 76 AlaIleArgGlyGluGluArgAspThrAspAlaAspAspGlySerAspPheAsnTyr 95
Db 471 CGCGTCCCGGGACGAGGAGGAGCGGCGTTCGACGACCTGGAGAACGAGTTCACTGG 530
QY 96 ProAlaSerGlyThrGluAspGlnLysGlnIleAlaAspArgMetArgSerTyrArg 115
Db 531 -----AGCGACAGACGACTCCAGTACTCGCCGAGTCCATGCTCCAGGCCAC 581
QY 116 MetAsnThrGlyGlySerGlyAsnValGly-----HisProLysTyr 129
Db 582 ATGAGTACGGCGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 635
QY 130 AspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIleProArgTyrVal 149
Db 636 -----ATCCCAAT-----GTT 647
QY 150 ProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis 169

```


Qy	865	ProLeuLeuValTyrCysIleLeuProAlaileCysLeuLeuThrGlyLysPheIleMet	884
Db	2805	CCGCTCTTTGGCGCTATTGCACATTCGCTGCCATCTGCTGTGCAGAGGAATTTATCACG	2864
Qy	885	ProGlutlleSerAsnLeuAlaSerlletTrpPheIleAlaLeuPheLeuSerillePheAla	904
Db	2865	::: ::: ::: ::: ::: ::: ::: :::	2924
Qy	905	ThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluUtpTrpArgAsnGlu	924
Db	2925	ACGAGCATCTTGAAATGAGATGGAGTGGTGTAGGCATCGATGACTGTTGTGGAGAACAAG	2984
Qy	925	GlnPheTrpValIlleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeu	944
Db	2985	CAGTTTTTGGGTCAVTTGGAGGCGGTCTCTTCAATCTTTGTGTGTCTCCAGGAGCTCCTC	3044
Qy	945	LysValLeuAlaGlyIleAspThrAsnPhetrValThrSerLysAlaAsnAspGluGlu	964
Db	3045	AAGTCAVAGTGTGTGTAGACCAGAGCTTCACTGTGACATCCAAGGGCGGAGACGACGAG	3104
Qy	965	GlyAspPheAlaGluLeuTyrMetPhelystrPrpThrThrLeuLeulleProProThrThr	984
Db	3105	::: ::: ::: ::: ::: ::: ::: :::	3161
Qy	985	IleleullelleAsnMetValGlyValValalaglyThrSerTyralalleAsnSerGly	1004
Db	3162	CTTGCTCTACTTGAATCTTCAATGGAGTGGTAGCTGGCATCTCCAATGGATCAACAACGGA	3221
Qy	1005	TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis	1024
Db	3222	TATGAATCATGGGGCCCCCTGTTCCGGAAGCTCTTCTTTGCATTTTGGGTGATGCTCCAT	3281
Qy	1025	LeuTyrProPheLeuLysGlyLeuMetGlyYargGlnAsnArghrProThrilValIle	1044
Db	3282	CTTTTACCCTGTTCTCAAGGGTCTGGTTGGGAGGCAGAACAGGACGCCAACGATTTGTCATT	3341
Qy	1045	ValTrpAlaValLeuLeuAlaSerillePheSerleuLeutrPvalArgValAspProphe	1064
Db	3342	GTCTGTGTCATCTCTGGCTTCGATCTTTCGCTGCTTTTGGGTCCGGATCGACCGGTC	3401
Qy	1065	ThrThrArgLeuAlaGlyProAsnIlleGlnThrCysGlyIleAsnCys	1080
Db	3402	CTTTCGGAAGATGATGTCCTCCCTGTTCCGAGAGGTGCTGTCGATTCC	3449

Schnable, Iowa State, then clones may be requested from www.zmdb.iastate.edu.

ORIGIN

Alignment Scores:			
Pred. No.:	0	Length:	3897
Score:	3806.00	Matches:	715
Percent Similarity:	77.24%	Conservative:	140
Best Local Similarity:	64.59%	Mismatches:	180
Query Match:	65.87%	Indels:	72
DB:	11	Gaps:	17
US-09-900-237A-30 (1-1080) x AY103655 (1-3897)			
QY	2	AspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIle	21
DB	222	GACGGCATCCCGGCCG---AAGCGCGCGGAGACAGACGGCAGGTGCCAGATT	278
QY	22	CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys	41
DB	279	TGCGCGACAGCTCGCGCTTGCCCCGCGGGACCCCTCGTGGCTGCAACAGATGC	338
QY	42	ArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCys	61
DB	339	GCCTTCCCCGTCTCCGGGACTGTCGAATACAGAGCCCGGGAGGACGCAGAACTGC	398
QY	62	LeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlu	81
DB	399	CCCAGTGCAGACTCGATACAAGCGCTCAAGGCTGCCAACGTGTGACCGTGACGAG	458
QY	82	GlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGlu	101
DB	459	GAGGAGACCGCGCTCGATGACCTGGACACACAGTTCAACTGGGACGGC---	506
QY	102	AspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyClySer	121
DB	507	CATGACTCGAGTCTGTGGCGGAGTCCATGCTC-----	539
QY	122	GlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSer	141
DB	540	-----TACGGCCACATGACTACGGCCGTGGAGGT-----GACCCT	575
QY	142	GlyGluIleProArgGlyTyr-----ValProSerValThrAsnSerGln	156
DB	576	AATGGCGCGCCACAAGCTTTCCAGCTCAACCCCAATGTTCCACTCTCCACCAACGGGCAA	635
QY	157	MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThr-	174
DB	636	ATGGTGGATGACATCCCA-----CCGAGCAGCAGCGGTGGTGGCTTCTTTTCATG	686
QY	175	---GlyAsnIleSerArgArgAlaProPheProTyrValAsn-----	187
DB	687	GTGTGTGGGGGAAGAGGATACATCCCTCTCTTATCGGATCCCGATTACCTGTGCAG	746
QY	188	---HisSerProAsnProSerArgGluPheSerGly---SerIleGlyAsnValalaTrp	205
DB	747	CCAGCTCTATGAGCCATCCCAAGGATCTTGTGCATATGGGTATGGTAGTGTGCTTGG	806
QY	206	LysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsn	225

Db	1017	TCAGCCGAATTAATCCCTACAGGATGATTATCGTTACCGGTGGTGGTTGGTTTC	1076	Db	2097	ATTTATGTGGGTACTGGATGTGTTTTCAGACGCGCAGCAGCTGTATGGTTATGATGCTCT	2156
Qy	289	PheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSer	308	Qy	649	lleIysAlaLysLysPro-----GlyPheLeuAla	658
Db	1077	TTCTTCCACTACCGAGTGATGCATCCGCGCAAGATGCATTTGGCATTTGGTGCATATCT	1136	Db	2157	---AAAAAGAGAACCCCATCAAGAACTTGCAACTGCTGCCCAAGTGGTGCCTCTCT	2213
Qy	309	ValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPhe	328	Qy	659	SerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLys	678
Db	1137	GTAATCTGTGAAATCTGGTTTCGGATGCTCTGGATTCTGTGATTCATCCCAAAGTGGT	1196	Db	2214	TCCTCTCGCAGCAGGAACAAGAACTACAAAACCAAAAGCAGGAGGAAGAAG	2273
Qy	329	ProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGlu	348	Qy	679	LysSer-----AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIle	696
Db	1197	CCATCGAGAGAGACTTACCTGCACCGTTTGTCACTAAGTTTGACAAAGGAGTCAA	1256	Db	2274	MAAGATTATTTTCAAGAAAGCAGAAAAACCATCTCTCTGCATATGCTTTGGGTGAAATT	2333
Qy	349	ProSerGlnLeuAlaValAspIlePheValSerThrValAspProLeuLysGluPro	368	Qy	697	GluGluGluValGluGluValaglyPheAspAspGluLysSerValLeuMetSerGlnMet	716
Db	1257	CCCTCTCAGCTTGCTCCAATCAGCTCTTTTGTGATGCTGATTCATCCCAAGGAACCT	1316	Db	2334	GATGAAGGTGCTCCAGGTGCT-----GATATCGAAGAGCCCGGAATCGTAATCAACAG	2387
Qy	369	ProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys	388	Qy	717	SerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyr	736
Db	1317	CCCTTGGTCACAGCAAACTGCTCTTCCATCCCTTCTGTGATTTATCCGGTTGAGAAG	1376	Db	2388	AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTGTGCGCATCAACACTTCTTGAGAAC	2447
Qy	389	ValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGlu	408	Qy	737	GlyGlyValProGlnSerSerThrProGlnSerLeuLeuLysGluAlaIleHisValIle	756
Db	1377	GTCTCTGCTATGTTCTGATCATGCTGCTGCAATGCTTACGTTTGAACATTTGCTGAA	1436	Db	2448	GGAGGGACCTCGAAGAGCGCAAGTCCAGCTCTCTTCTGAAGGAAGCTATATGTTATC	2507
Qy	409	ThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArg	428	Qy	757	SerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySer	776
Db	1437	ACATCTGAATTTGCCAAGAAATGGGTTCTCTTTCAGCAAAAGTTTAAATATCGAGCTCGT	1496	Db	2508	AGCTCGGCTTACGAAGACAAGACCAGCTGGGGAAGAGATTGGCTGGATTACGATGCG	2567
Qy	429	AlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSer	448	Qy	777	ValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyr	796
Db	1497	GCTCTGAGTGGTACTTCCAAAGAGATAGACTACCTGAAAGACAAAGTTTCTGCTTCA	1556	Db	2568	ATCACAGAGGATATCTTGACTGATTTAAGATGCACTGCCATGGCTGGCGGCTATTAC	2627
Qy	449	PheValLysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsn	468	Qy	797	CysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeu	816
Db	1557	TTTGTAGGAGAGAGGCGCATGAAGAGAGATACAGAGAAATCAAGGTAAGGATCAAT	1616	Db	2628	TGCACTCCCGAAGCGGCTGCATTTCAAAGGTTCTGCGCTCTGAAACCTTTCCGACCGCTCT	2687
Qy	469	AlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThr	488	Qy	817	AsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeu-PheSerArgHisCysPr	836
Db	1617	GCCTTGGTTGCAAAAGCCCAAAAGGTTCTCGAAGAGGATGACAAATGCAAGATGGAAGC	1676	Db	2688	CACCAAGTCTTCGCTGGGCCCTTGGTCCGTCGAAATTTCTTTCAGNAANCAATGGNC	2747
Qy	489	ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis	508	Qy	836	oLeuTrpTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnTh	856
Db	1677	CCCTGCGCTTGGAAACAACGTACCGCATCTCTGGAATGATTACAGTATTCCTTGCCCAA	1736	Db	2748	ACTTTGGTACGATACGCGCGCGGCTAAAAATTCCTGGAAGAGTTTCTTATATCACTC	2807
Qy	509	SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu	528	Qy	856	rThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCy	876
Db	1737	AGTGGCGTCTGTGATGTGAAGAAATGAGTTGCTCGCTCGCTGTTTATGTCTCGAGAA	1796	Db	2808	CATCGTTTATCCCTGACGCTCATCTCTCTGCGTTACTGTACCTTCCTGCGCACTCG	2867
Qy	529	LysArgProGlyPheGlnHisIleLysLysAlaGlyAlaMetAsnAlaLeuValArgVal	548	Qy	876	sLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIle	896
Db	1797	AAAGAGCCAGGTATAACCATCACAAAGAGGCTGGTGCCATGAATGCACTGTGCTGTC	1856	Db	2868	CCTGCTCACGGGAAGTTTATCACACAGAGCTTACCATGTCCGCGCATCGACAGCTTACCGT	2927
Qy	549	SerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsn	568	Qy	896	eAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlu	916
Db	1857	TCGTGCTTCTTCAATGCTGCATACCTATTGAACTTGACTGTGATCACTATCATCAAC	1916	Db	2928	GGCACTTTTTCATCTGCATCTCCGTGACCGGCTCTCTGCGTTACTGTACCTTCCTGCGCATCG	2987
Qy	569	AsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGln	588	Qy	916	ylleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisIle	936
Db	1917	AATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGATGCTTTTGGTGGGAAGAAA	1976	Db	2988	CATCGACGACTGGTGAGGAACGACGAGTTCTGGGTCATCGGAGGCGTTTCGGCGCATCT	3047
Qy	589	ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAla	608	Qy	936	uPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrVa	956
Db	1977	GTGTGCTATGTACAGTTCCCTCAGAGGTTTGATGTTGATTTGACACAAATATGATGATACGCT	2036	Db	3048	GTTCCGCGTTCACAGGCGCTGCTGAAGGTGTTCGCGCGCATCGACACGAGCTTACCGT	3107
Qy	609	AsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyPro	628	Qy	956	lThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpTh	976
Db	2037	AAACGGAACGTTGCTCTTTTGTGATCAACATGAAAGGTTTGACCGGTATTCAAGGACCC	2096	Db	3108	GACGTCGAAGGCGGCGGACGACGAG---GAGTTTCTCGAGCTGTACACGTTCAAGTGGAC	3164
Qy	629	ValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProPro	648	Qy	976	rThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlu	996
Db				Db	3165	CACCTTGTGATACCCCCGACACGCTCTCTCTGCTGAACTTCTATCGGGGTGGTGGCCGG	3224

```

Qy 996 yThrSerTyrAlaIleAanSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPh 1016
Db 3225 GATCTCGAACGGATCAACAACGGGTACAGTCGTGGGGCCCCCGTTCGGGAGACTCTT 3284
Qy 1016 ePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgG1 1036
Db 3285 CTTGGCGTTCTGGGTGATCGTCCACCTGTACCGGTCTCTCAAGGGTCGTGGTGGGAGGCA 3344
Qy 1036 nAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLe 1056
Db 3345 GAACAGGAGCGCGAGATCGTCATCGTCTGTGGTCCATCTGTGCTGCTCGATCTTCTCGCT 3404
Qy 1056 uLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAanIleGlnThrCy 1076
Db 3405 CCTGTGGGTCTCGGTGACCGCTTCTCTCGCCAAAGAGCGACGGCCGCTCCTGGAGGAGTG 3464
Qy 1076 sGlyIleAanCys 1080
Db 3465 TGGCTGGACTGC 3477

RESULT 8
AY104236
LOCUS Zea mays PC0121439 mRNA sequence.
DEFINITION Zea mays PC0121439 mRNA sequence.
ACCESSION AY104236
VERSION AY104236.1 GI:21207314
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2872)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2872)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
source
1..2872
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="maizeDB:637064"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 8,76e-309 Length: 2872
Score: 3525.00 Matches: 646
Percent Similarity: 88.42% Conservative: 87
Best Local Similarity: 77.93% Mismatches: 85
Query Match: 61.01% Indels: 11

```

```

DB: 11 Gaps: 6
US-09-900-237A-30 (1-1080) x AY104236 (1-2872)
Qy 258 ArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAanProTyrArgMet 277
Db 16 CGGCTACCTCTTAAGTCGATAGTTCCGATATCTCCAAAGAGCTTAACCTTATCGGATC 75
Qy 278 ValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnPro 297
Db 76 GTGATTGTTCCTCGCGCTTATCATCTCTATGTTTCTTCTCAATATCGTATAACTCATCCA 135
Qy 298 ValArgAsnAlaTyrProLeuTrpLeuSerValIleCysGluIleTrpPheAlaLeu 317
Db 136 GTGGAAGATGCTTATGGGTGTGTCTGTATCTGTATTTGTGAAGTTTGGTTTGGCTTG 195
Qy 318 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAanArgGluThrTyrLeuAsp 337
Db 196 TCTTGGCTTCTAGATCAGTTCCTCAAGTGGTATCTTATCAACCGTGAACCTTACTCGAT 255
Qy 338 ArgLeuAlaLeuArgTyrAspArgGluGluProSerGlnLeu-AlaAlaValAspI1 357
Db 256 AGACTTGCTTACATATGATAGGAGGGTGAGCCATCCAGTTGGGCTCCAAATCGATGT 315
Qy 357 ePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLe 377
Db 316 CTTTGTGTAGTACAGTGGATCCACTTAAGGAACCTCTCTAATTACTGGCAACACTGTCT 375
Qy 377 userIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspG1 397
Db 376 GTCCATCTTCTGCTGGATTACCTCTGTGACAAAGTATCATGTTATGTTTCTGATGACGG 435
Qy 397 yAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpVa 417
Db 436 TTCAGTATGTTGACTTTTGAAGCGCTATCTGAAACCGCAGAGTTTGCAGAAATCGCT 495
Qy 417 lProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLy 437
Db 496 TCCCTTTTGAAGAAACACAAATATTGAACCTTAGGGCTCCAGAGTTTACTTTGCTCGAAA 555
Qy 437 sIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLy 457
Db 556 GATAGATTACTTAAGGACCAAAATACAACTCTTTTGTGNAAGAAAGCGGGCTATGAA 615
Qy 457 sArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysVa 477
Db 616 GAGGGAGTGTGAAGAGTTCAAAGTACGATCGCTTGTTCGAAAAGCGCAAAAAAT 675
Qy 477 lProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAs 497
Db 676 ACCTGAGGAGGGCTGGACCATGGCTGATGGCACTCCCTGGCTGGGAATAACCCCTAGAGA 735
Qy 497 pHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAs 517
Db 736 TCATCCMGGAATGATCCAGTATTTCTTGGGCCACAGTGTGGGCTTGACACGATGGGAA 795
Qy 517 nGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLy 537
Db 796 TGAGTTGCCACGGCTTGTATGTTTCTCGTGAAGAGGGCCAGGCTTCAGCACCAACAA 855
Qy 537 sLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTy 557
Db 856 GAAGGCTGGTGGCATGAATGCTTGTGATTCGCGTATCAGCTGTCTGACGAAATGGTGTGA 915
Qy 557 rMetLeuAsnLeuAspCysAspHisTyrIleAanSerLysAlaValArgGluAlaMe 577
Db 916 TCTTCTTAATGTGATGTGATCACTACTTCAATAGCAGCAAGAGCTTTAGAGAGGCTAT 975
Qy 577 tCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnAr 597
Db 976 GTGTTTTCATGATGGATCCAGACTTAGGAAGGAAAACTTGCTATGTTTCCAGTTTCCACAAAG 1035
Qy 597 gPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspI1 617

```



```
Score: 2507.00 Matches: 462
Percent Similarity: 84.44% Conservative: 70
Best Local Similarity: 73.33% Mismatches: 84
Query Match: 43.39% Indels: 14
DB: 14 Gaps: 5

US-09-900-237A-30 (1-1080) x CD726831 (1-1874)

QY 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrIleuAspArgLeu 339
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 2 GTGCTGGACGATTCCTTAATGGTTACCTGCTGCAGGGACACATTTATTGACAGATTA 61

QY 340 AlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheVal 359
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 62 TCAGCCAGGTTTGAAGAGAGAGGAGAACCTTCCAGCTTGCTGCTGGATTTTGTG 121

QY 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 122 AGTACCGTTGATCCGTTGAAGAACCTCTCTAATCACTGCGAATACCGTGCTTCAATC 181

QY 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 182 CTTCGTGGACTATTCCTGCTAATAGTCTCTGCTATGTGTCGATGACGGTGACGCT 241

QY 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaAlaArgLysTrpValProPhe 419
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 242 ATGCTCACATTGGAATCTAGTTGAACAGCTGACTTTGCAAGGAATGGTTCATTC 301

QY 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 302 TGCATAAAATTTCTCCATTGAACCTCGAGCTCTCGAGTTTACTTCTCCAGAAAATAGAC 361

QY 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAtcAlaMetLysArgGlu 459
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 362 TATTTGAAGAATAAGTACAACTTCTTTGTAAGGAACGTAGAGCTATGAAGAGAGCT 421

QY 460 TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 422 TATGAAGAGTTTAAAGTACGAGTCAATGCTTTGGTGGCAAGGGCCCAAAACACCTGAT 481

QY 480 GluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 499
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 482 GAAGGCTGGTCCATGCAAGATGGAACAGCTTTGGCCAGGAACAATCCACGTCATCACCT 541

QY 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 542 GGAATGATCAGGTTTTCTTGGAAATACCGCGCCCATGACGTAGAGGGAATGAACCT 601

QY 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 539
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 602 CCTCGACTAGTCTACGCTCCAGAGAGAAGAGGCTGGTTACCAGCATCACAAAAGAGCT 661

QY 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 662 GGTCTGAAATAGTCTGGTAAGAGTATCTCAGTCTCTCAAAATGCCCCCTTCATCTT 721

QY 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 722 AACCTTGACTGTGATCATCTAGTTAAACAATAGTCAAGCTATACGTAGGCAATGTGTTTC 781

QY 580 LeuMetAspProAsnLeuLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 782 TTGATGACCCACAAAGTTGGCCGAGATGATGCTATGTTCAAGTTTCCTCAGAGGTTTGTAT 841

QY 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 842 GGCATTGATCGAGTGATCGATATGCAATGCCAACCGGTTTCTTCGATGTGAACATG 901

QY 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 902 AAAGATTTGGATGGCATCCAGGACCAAGTTTATGTTGTACAGGTTGTGTTTCAATAGA 961

QY 640 ThrAlaIleTyrGlyTyrGluProProIleLys-----AlaLysLysProGlyPhe 656
```

```
Db 962 CAGGCACCTTTATGGCTACGACCTCTCTCCATGCCCTAGCTTATCTAAGACATCCTCATCA 1021
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
QY 657 LeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysLysLysArgSerSerAsp 676
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1022 TCTTGTCTCTGTGTGGCTGTGCTCTTGTCTCTTCCCAAGAAGATCTCAAAAGAT 1081

QY 677 Lys-----LysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1082 CCGACTGAGATTCAGAGAGATGCAAAAAGAGAGAGCTTGATGCTGCAATCTTTAACCTC 1141

QY 694 GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAsp---GluLysSerValLeu 712
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1142 AGGGAATAGAT-----AATTATGATGAGTATGAGAGATCAATGCTG 1183

QY 713 MetSerGlnMetSerLeuLysArgPheGlyGlnSerAlaAlaPheValAlaSerThr 732
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1184 ATTTCTCAACTGAGCTTTGAGAAAATTTGCGATTTGCTGCTGTTTATCGAGTCTAGC 1243

QY 733 LeuMetGluTyrGlyValProGlnSerSerThrProGluSerLeuLysGluAla 752
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1244 CTAATGGAATAATGGCGAGTTTGGGAATCTGCAATCCCTCGACTTTGATCAAGGAAGCA 1303

QY 753 IleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTrp 772
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1304 ATTCATGTCATAGCTGTGGTTTATGAGAGAAGTCCACTTGGGGGAAAAGAGATTGGTGG 1363

QY 773 IleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrp 792
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1364 ATATATGGGTGAGTACGAGTATCTTAACGGGTTTCAAGATGTCATTTGCCAGGGTGG 1423

QY 793 ArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeu 812
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1424 AGGTCCATCTACTGATGCCATTTAGGCCAGCAITTCAAAGGGTCCGCAACCAATTAACCTT 1483

QY 813 SerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSer 832
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1484 TCTGATCGTCTCCCAAGTTCTTCGATGGGACATTTGGATCTGTGTGAGATTTTCTTAGC 1543

QY 833 ArgHisCysProLeuTyrTrpGlyTyr---GlyGlyArgLeuLysPheLeuGluArgPhe 851
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1544 AGACACTGTCCATTTATGATGGATTTTCAGCGCGCGCCCTCAATGGCTCCAAAGAATG 1603

QY 852 AlaTyrIleAsnThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIle 871
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1604 GCITTACATAAAACACCATTTGCTATCCCTTTCACCTCGCTCCCTCTTGTGTACTGCTCA 1663

QY 872 LeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAla 891
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1664 TTGCTCGCATCTGCTGCTCACAGGAAGTTTCATTCATTCACAGCTCTCGAACCTGACA 1723

QY 892 SerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArg 911
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1724 AGTACCTTTTCTCGGTCTCTCTTGTCCATCATTTCTCAGAGTGTCTCGAGCTGGT 1783

QY 912 TrpSerGlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGly 931
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1784 TGGAGTGGTGTAGCATCGAAGATATATGGCGTAACGAGCAATTTCTGGGTAATCGGAGGC 1843

QY 932 IleSerAlaHisLeuPheAlaValPheGln 941
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1844 GTCCCGGCACATCTCTTTGGCGTCTTCCAA 1873

RESULT 10
CK272603 916 bp mRNA linear EST 12-DEC-2003
LOCUS EST718681 potato abiotic stress cdna library Solanum tuberosum cDNA
DEFINITION clone POD624 5' end, mRNA sequence.
ACCESSION CK272603
VERSION CK272603.1 GI:39829581
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE AUTHORS TITLE JOURNAL COMMENT

1 (bases 1 to 916)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: ESR718682
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: Arr TAG GTG ACA CTA TAG.

FEATURES

Location/Qualifiers
1..916
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOAD624"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed with withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.: 2,69e-132 Length: 916
Score: 1575.00 Matches: 283
Percent Similarity: 99.34% Conservative: 20
Best Local Similarity: 92.79% Mismatches: 0
Query Match: 27.26% Indels: 0
DB: 14 Gaps: 0

US-09-900-237A-30 (1-1080) x CK272603 (1-916)

QY 738 GlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaHisValIleSer 757
Db 1 GGTGTTCTCAATCGGTACCCGGAGACCTTTTGAAGAGCGTATTCATGATCAGT 60
QY 758 CysGlyTyrGluAspLysSerGluThrGlyThrGluLeuGlyTrpIleTyrGlySerVal 777
Db 61 TGTGGTTATCAAGATAAATCAGATGGGAACTGAGATGGATGATGTTGGTTCGTC 120
QY 778 ThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCys 797
Db 121 ACAGAGGATATCTTCTTGGATTTAAGATGATCGCCGTTGGTGGCGATCTATTACTGT 180
QY 798 MetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsn 817

Db 181 ATGCCCAAGAGAGACCCGCCCTTCAAAGGGTCAGTCTCTATTAAATCTTTTCAGATCGTCTGAAC 240
QY 818 GlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeu 837
Db 241 CAAGTGCCTTCGATGGCGCTTTAGGTCAGTGGAAATCTTTTCAGTAGGATTTGCTCTATA 300
QY 838 TrpTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThr 857
Db 301 TGGTATGATACATAATGGACGGTTGAAGTGGTGGAGAGATTGCTTATGTCACACACC 360
QY 858 IleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeu 877
Db 361 ATTTATCCAATCATTCCATCTCTATATATCTGATCTCCAGCATCTATCTCTA 420
QY 878 LeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAla 897
Db 421 CTTACTGGGAAATTCATTATCTCTGATGATGATGATGATGATGATGATGATGATGAT 480
QY 898 LeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIle 917
Db 481 CTCTTTCTTTCCATTTTCTGCTACTGCTATCTGGAGATGAGATGGAGTGGTGTGAAT 540
QY 918 AspGluTrpTrpArgAsnGluInPheTrpValIleGlyGlyIleSerAlaHisLeuPhe 937
Db 541 GATGAATGGTGGAGAAATGAACAGTTTGGGGTCATTGGTGGTGTGTCACCTGCTTT 600
QY 938 AlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 957
Db 601 GCGCTTCCAAAGGGTGTCTCAAGTGTGTGCTGTTATGATACCAACCTTACTGTGACA 660
QY 958 SerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThr 977
Db 661 TCCAAGGCATCAGATGAAGATGGGACTTTGGGAACTCTACTTGTTCAAATGGACAAC 720
QY 978 LeuLeuIleProThrThrIleLeuIleLeuIleAsnMetValGlyValValAlaGlyThr 997
Db 721 CTCTTTATACCCCCCACTACTCTCTCATTTGTAACCTGCTAGAGTGTGGCGGCATC 780
QY 998 SerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePhe 1017
Db 781 TCATATGCCATCAACAGTGGTACCACATCATGGGGTCCCTCTTTGGTAAATATTCTTT 840
QY 1018 AlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsn 1037
Db 841 GCCTTTCTGGGTGATCGTTCACCTTTACCCCTTCCCAAAGGTCCTCATGGGTGCTCAGAAC 900
QY 1038 ArgThrProThrIle 1042
Db 901 CGGACACCCCACTATC 915

RESULT 11

CF513822

LOCUS

DEFINITION

CF513822

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CF513822 924 bp mRNA linear EST 09-SEP-2003
Cabud0007_IIIF_B10 Vitis vinifera cv. cabernet sauvignon (Clone 8)
Bud - CABUD Vitis vinifera cDNA clone Cabud0007_IIIF_B10 5', mRNA
sequence.

CF513822.1 GI:34545590
EST.
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 924)
Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Jones, K. and
Cook, D.

Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
Unpublished (2003)
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seq primer: ACGGTACCGACATATGCC.

FEATURES source

1..924
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon (Clone 8)"
 /db_xref="taxon:29760"
 /clone="Cabud0007_IIIF_B10"
 /sex="Hermaphrodite"
 /dev_stage="Pre-bloom (10-11 days before bloom)"
 /lab_host="DH5alpha"
 /clone_lib="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud- CABUD"
 /note="Organ: Bud; Vector: pDNR; Site: 1: SfiI; Site 2: SfiI; CABUD is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGTGATCAACGACGAGTGGCCATTACGGCGGG-3' and
 5'-ATTCTAGAGCGGCGGCCACATG-dT(30)/NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Alignment Scores:
 Pred. No.: 2,7e-129 Length: 924
 Score: 1542.00 Matches: 277
 Percent Similarity: 98.37% Conservative: 25
 Best Local Similarity: 90.23% Mismatches: 4
 Query Match: 26.69% Indels: 1
 DB: 14 Gaps: 0
 US-09-900-237a-30 (1-1080) x CF513822 (1-924)

QY 775 GlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSer 794
 DB 2 GGGTCTGTACGGAAGATATCTTACAGATTCAAGATGTCATGCTCGGGTGGAGATCT 61
 QY 795 ValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAsp 814
 DB 62 ATTACTGCATGCCCAAGCGCCGCCCTTCAAGGGGTCTGCCCTATTAATCTTTTCAGAT 121
 QY 815 ArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHis 834
 DB 122 CGTCTGAACCAAGTCTTCGATGGCCCTCGGTCTGTGGAATCTTTTCAGTCGGCAC 181
 QY 835 CysProLeuTyrPyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
 DB 182 TGTCTATCTCGTATGTTATGTTGGTGAAGGCTAAATGGCTTGAGATTCGCATATGTA 241
 QY 855 AsnThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
 DB 242 AACACCAACCATTTATCAATACCGCATCCCTCTCTGTTTACTGCATCTCGCCGCT 301
 QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyr 894
 DB 302 GTTTGTTTGTCTACTGGAAGTTTCATTATCCAGATCAGTACATGATGCAAGTATTGG 361
 QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGly 914
 DB 362 TTTATATCCCTCTCTCTTCATCTTTGCTACCGGTATATTGGAGATGAGATGGAGTGGT 421
 QY 915 ValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyIleSerAla 934
 DB 422 GTTGGATCATGATGAATGGTGAAGAAATGAGCAGTTTGGGTATTGGTGGCTCTGCT 481

QY 935 HisLeuPheAlaValPheGlnGlyIleuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
 DB 482 CATCTTTTGTCTGTTTTCGAAGCCCTGCTGAAAGTACTTGTCTGAATTCATACCACTTC 541
 QY 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
 DB 542 ACTGTCACTCCCAAGCTTCGGATGAAGACGGGGATTTTGTGGAATCTACATGTTTAA 601
 QY 975 TrpThrThrLeuLeuIleProThrThrIleLeuIleIleAsnMetValGlyValVal 994
 DB 602 TGGACAACCTCTTCTCATCCACCCCAACAGCTCCTCATATCAATCACTGTGGGTGTGT 661
 QY 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLys 1014
 DB 662 GCAGGGATCTCTCATGTCTCAACAGTGGCTACCAAGTCAATGGGGTCCACTCTTTGGCAAG 721
 QY 1015 LeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
 DB 722 CTCCTCTTGTCTTGGTGTGGTGTATGCTTCACTCTATCCCTTCTCTCAAGGTCTGATGGA 781
 QY 1035 ArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIlePhe 1054
 DB 782 CGCCAGAAACCGAACCACTTATGTTGTCTGTCGATTCCTCTCTTCCATTTTC 841
 QY 1055 SerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGly-ProAsnIleGly 1074
 DB 842 TCGTGTGTTATGGTGCAGAAATTGATCCCTTCCACTAGATCAGTGGCCCCAGATGTTGA 901
 QY 1074 nThrCysGlyIleAsnCys 1080
 DB 902 GCAGTGTGGTATCAACTGC 920

RESULT 12
 CK271294
 LOCUS
 DEFINITION
 EST717372 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POACY34 5' end, mRNA sequence.
 CK271294
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.
 Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)

Other ESTs: EST717373
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr. Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES source

1..977
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="kennebec"
 /db_xref="taxon:4113"
 /clone="POACY34"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 °C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 °C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:	3 22e-128	Length:	977
Pred. No.:	1530.50	Matches:	285
Score:	95.06%	Conservative:	23
Percent Similarity:	87.96%	Mismatches:	13
Best Local Similarity:	26.49%	Indels:	2
Query Match:	14	Gaps:	2
DR:			

US-09-900-237A-30 (1-1080) x CK271294 (1-977)

645	Qy	TyrGluProProlLeLysAla-----LysLysProGlyPheLeuAlaSerLeuCysGly	566
		6 TATGAACCTCCAAATTAGACCAAGCATAGAAGCAGGGTTCCCTCTTCTCTCGTCCGT	65
663	Qy	Gly---LysLysLysAlaSerLysSerLysArgSerSerAspLysLysLysSerAsn	681
		66 GGATCAGAAAGAGAGGGTTCTAATTCAGATATAAAGAGCTCAGACAAGAAGAACTAGT	125
682	Qy	LysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGlyValGlu	701
		126 AAGAAATGTTGATCCCACTGTCCTAATTCATCTGGAGGATATAGAGGAGGAGTTCGA	185
702	Qy	GlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArg	721
		186 GGTGCTGGATTTGATGATGAGAAGTCACCTTCATGTCACAAATGAGCTGGAGAAGAGA	245
722	Qy	PheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValProGln	741
		246 TTTGGGCAATCGGCTGTTTTGTTGCTTCAACACATCATGGAGAATGGTGGTGTCTCTCAA	305
742	Qy	SerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGlu	761
		306 TCGGCTACCCACAGACACCCCTTTTGAAGAGGGCTATTTCATGTTATCAGTTGTGGTATTATGAA	365
762	Qy	AspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAspIle	781
		366 GATAAATCAGAAATGGGAACACTGAGATTGATGGATCTATCGTTCCGTACAGAGGATATT	425
782	Qy	LeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLysArg	801
		426 CTTACTGGATTTAAGATGTCATGCCGTGGTGGCGATCTATATCTGTATGCCCAAGAGA	485
802	Qy	ProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg	821
		486 CCCGCGCTTCAAAAGGGTCAGTCCCTAATAATCTTTTCAGATCGTCTGAACCAAGAGCTTCGA	545
822	Qy	TrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGlyTyr	841
		546 TGGGCTTTAGGGTCAGTGGAAATCTTTTTCAGTAGGACATGTTCCTATATGTTATGGATAC	605
842	Qy	GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeu	861
		606 AATGACACGCTTCAAGTGGTTGGAGATTGCTTATGTCAACACCAACCAATTTATCCAAATC	665

Qy	862	ThrSerLeuProIleuValTyrCysIleLeuProAlaIleCysIleLeuThrGlyLys	881
Db	666	ACTTCACATCCACTCTTTATATACTCATGCATCTCCAGACTATCTGTACTACTCGGAAA	725
Qy	882	PheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSer	901
Db	726	TTCATTATCCCTCAGATTAGTAACCTTGCTAGCATCTGGTTTTATATCCCTCTTCTTTC	785
Qy	902	IlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrpTrp	921
Db	786	ATTTCGCTACTGCTATCTCGAGATGAGATGAGTGTGTGGAATTGATGAATGTGTGG	845
Qy	922	ArgAsnGlnGlutPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPheGln	941
Db	846	AGAAATGAACAGATTTGGGGTCATTTGGTGGTGTGCACCTCACCTGTTGGCTTCCAA	905
Qy	942	GlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsn	961
Db	906	GGGTTCCTCAAGTGTCTGTGTAATTGATACCACCTTTACTGTGCATCCAGGCAACA	965
Qy	962	AspGluGluGly	965
Db	966	GATGAAGATGGG	977
RESULT 13			
CK274934			
LOCUS			
DEFINITION			
EST721012 potato abiotic stress cdna library Solanum tuberosum cDNA clone POADK31 5' end, mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Solanum tuberosum (potato)			
Solanum tuberosum			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.			
1 (bases 1 to 958)			
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.			
Generation of ESTs from abiotic stressed potato tissue			
Unpublished (2003)			
Other ESTs: EST721013			
Contact: Robin Buell			
The Institute for Genomic Research			
9712 Medical Center Dr, Rockville, MD 20850, USA			
Email: potato-array@tigr.org			
Clones can be requested from TIGR via potatoetigr.org			
Seq primer: ATT TAGT GTG ACA CTA TAG.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

FEATURES

```

1. .958
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POADK31"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TorA"
/clone_lib="potato abiotic stress cDNA library"
/note="vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

```

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

```

Pred. No.:      8,9e-126      Length:      958
Score:          1503.50      Matches:      281
Percent Similarity: 95.00%      Conservative: 23
Best Local Similarity: 87.81%      Mismatches: 12
Query Match:      26.02%      Indels:      4
DB:              14          Gaps:      2

US-09-900-237A-30 (1-1080) x CK274934 (1-958)

QY 633 ThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla--- 651
Db 1 ACTGGATGCTTCAATAGAACACCTTATATGGTTATGACCTCCCAATTAAGCCAAAG 60
QY 652 ---LysLysProGlyPheLeuAlaSerLeuCysGlyGly---LysLysLysAlaSerLys 669
Db 61 CATAAGAGGCGAGGCTTCTCTCTCCGCTTCGGTGGATCAAGAAAGAGGTTCTTAAT 120
QY 670 SerLysLysArgSerSerAspLysLysSerLysHisValAspSerSerValPro 689
Db 121 TCAAGTAAAGGAGGCTCAGACAAAGAAATCTAGTAAAGATTTGATCCCACTGTGCCA 180
QY 690 ValPheAsnLeuGluAspIleGluGluValGluGluValGluGluPheAspAspGluLys 709
Db 181 ATATTCAATCTGGAGGATATAGAGAGGAGTTGAAGGTGCTGGATTTGATGATGAGAAG 240
QY 710 SerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheVal 729
Db 241 TCACCTTCTCATGTACAAATGAGCCTGGAGAAGAGATTTGGCAATCGGCTGTTTGT 300
QY 730 AlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeu 749
Db 301 GCITTCACACACTCATGGAGATGGTGGTGTCTCCTCAATCGGCTACCCAGAGACCCCTTTG 360
QY 750 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGlu 769
Db 361 AAGAGGCTATTCAATGTTATCAGTTGTTATGAGATAATCAGAAATGGAATGGGAACTGAG 420
QY 770 IleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAla 789
Db 421 ATTGGATGGATCATGGTTCGGTCACAGAGGATATTCTTACTGGATTTAAGATGCATGCC 480
QY 790 ArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaPro 809
Db 481 CGTGGTTGGCGATCTATATAGCTATGATCCCAAGAGACCCGCTTCAAGGCTCAGCTCCT 540
QY 810 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIle 829
Db 541 ATTAATCTTTCAGATGCTGTAACCAAGTCTTCGATGGGCTTTAGGGTCAGTGGAAATT 600
QY 830 LeuPheSerArgHisCysProLeuTyrTyrGlyTyrGlyGlyArgLeuLysPheLeuGlu 849
Db 601 CTTTTAGTAGGCATTTCTCTATATGATGATGATACATGACGAGGTTGAAGTGGTGGAG 660
QY 850 ArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyr 869
Db 661 AGATTGCTTATGTCACACACCACTTATCCCAATCACTTCCATTCCTTATATATAC 720
QY 870 CysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsn 889
Db 721 TGCATGCTCCAGCTATCTGCTACTTACTGGGAAATTCATTATCCCTCAGATTAGTAAC 780
QY 890 LeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGlu 909

```

```

Db 781 CTTCGATGATCTGGTTATATATCCCTCTTCTTCCATTTCTGCTACTGTTCTGGAG 840
QY 910 MetArgTrpSerGlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIle 929
Db 841 ATGAGATGAGTGGTGGTGGATTCATGATGATGGTGGAGAAATGACAGATTTGGGTCATT 900
QY 930 -GlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla 948
Db 901 GGGTGGTGTGTCAGCTCACCTGTTTGGCGCTCTTCCAAGGGTTGCTCAAAGTCTTGTCT 958

```

RESULT 14

BG321258

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

FEATURES

source

```
Db 123 GATGGTATTGATGAAATKACCGKATCCAAACAGGACACATGTGTTTTTCGATTTAAAC 182
Qy 619 LeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsn 638
Db 183 TTGAGAGCKTTGATGCAATTCAGGGCCAGTTTATGKGAACCTGGTGTGTGTTTAAAC 242
Qy 639 ArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAla 658
Db 243 AGAAGCGCTTATATGTTATKAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTCT 302
Qy 659 SerLeuCysGlyGlyLysLysAla-SerLysSerLysLysArgSerSerAspLysIle 678
Db 303 TCGCTTTTGGGGAGAGAAAAAGACGCTCAAAATCTAAGAAG---AGCTCGAAAGAA 359
Qy 678 sLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluCl 698
Db 360 GAAGTCAATAGACACAGCAGACAGTCTCTGTACAGTATTATCTCGAAGATATAGAGGA 419
Qy 698 uGlyValGluGlyAlaGlyPheAspGluLysSerValLeuMetSerGlnMetSerLe 718
Db 420 AGGGAATTTAAGGTTCTCAGTTTGTATGAGAAATCGCTGATATATGTTCTCAATGAGC 479
Qy 718 uGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyCl 738
Db 480 KGAGAGAGATTTGGCCAGTCAGTCTTTTGTAGCCCTACTCTCATGGAATATGGCGG 539
Qy 738 yValProGlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCy 758
Db 540 TGTTCACAAATCTGTACTCCGAGTCTCTCTGAAAGAGCBATTCATGTGCATCAGCTG 599
Qy 758 sGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValTh 778
Db 600 TKGCTATTAGAGCAAAACAGTGGGGAACAGTGGGGAACAGTGGGGAACAGTGGGGA 659
Qy 778 rGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMe 798
Db 660 AGAAGACATCTCACAGGATTCAGATGATGATGCBGAGGCTGGCGATCAATCTACTGCAT 719
Qy 798 tProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGl 818
Db 720 GCTTAARCSACCAGCTTTCAAGGGAATCTGCTCCYATCAACCTTTTCGGATCGTTGA 779
Qy 818 nValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTr 838
Db 780 AGTGCYTCGGTGGCTCTTGGKTCATTAATKAAATCCCTTTTCAGCAGCATTTGCCAT 839
Qy 838 pTyrGlyTyrGlyArgLysPheLeuGluArgPheAlaTyrIleAsnThrThrI1 858
Db 840 GTATGCTATTGAGCGCGCTTAATTCCTGGAGARATTGGCTTATATCAACACAACAT 899
Qy 858 eTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLe 878
Db 900 TTATCCACTCACATCAATCCGCBCTCYGTACTGCATATTCGCCAGCT-TGTCTTCT 958
Qy 878 uThrGlyLysPheIle 883
Db 959 CACTGGGAAGTTTCATC 974
```

RESULT 15

```
CD439980 953 bp mRNA linear EST 03-JUN-2003
LOCUS EL01N0531D05.b EndospERM_5 Zea mays cDNA, mRNA sequence.
DEFINITION CD439980
ACCESSION CD439980
VERSION GI:31355623
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 953)
AUTHORS Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
```

```
Messing,J.
Sequencing of the maize endospERM ESTs
Unpublished (2002)
Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3
Location/Qualifiers
1. 953
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="EndospERM of 7-23DAP"
/clone_lib="EndospERM_5"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 4,19e-119 Length: 953
Score: 1430.00 Matches: 257
Percent Similarity: 91.64% Conservative: 28
Best Local Similarity: 82.64% Mismatches: 25
Query Match: 24.75% Indels: 1
DB: 14 Gaps: 0
```

US-09-900-237A-30 (1-1080) x CD439980 (1-953)

```
Qy 728 PheValAlaSerThrLeuMet-GluTyrGlyValProGlnSerSerThrProGluSe 747
Db 21 TTCATGTGATCCACCTTTATGACACACCGTCGCATACCACCTTCAACAAACCCAGCTTC 80
Qy 747 rLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGl 767
Db 81 TCTTACTAAGGAAGCTATCCATGTGCATGTGATGATGAGGACAAACTGAATGGGG 140
Qy 767 yThrGluIleGlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMe 787
Db 141 AAAAGAGATTTGGCTGATCTATGTTTCAAGGAGATATTCCTGACGGGTTTAAAT 200
Qy 787 thisAlaArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySe 807
Db 201 GCATGCAAGGGCTGGCAATCAATCTACTGTCATGCCACCCAGACCTTGTTCAGGGTTC 260
Qy 807 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerVa 827
Db 261 TGCACCAATCAATCTTCCGATCGTCTTAATCAGGTGCTCCGTGGGCTCTTCGGGTCACT 320
Qy 827 lGluIleLeuPheSerArgHisCysProLeuTyrGlyTyrGlyArgLysLysPh 847
Db 321 GGAATTTCTCTTAGTAGACATTTCTCTATCTGGTATGGTTTCAATGGACGATTTAGACT 380
Qy 847 eLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLe 867
Db 381 TTTGGAGAGGCTGGCTTACATCAACACTATTGATATATCCAAATCACAATCCGCTTAT 440
Qy 867 uValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluI1 887
Db 441 TGCCTATTGTGTCTTCCGCTATCTGCTCTTACCAATAAATTTATCATCTCTCTGAGAT 500
Qy 887 eSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyI1 907
Db 501 TAGCAATATTGCTGGGATGTTCTTCATCTCTTTTCGCCCTCCATTTTTCACCTGCTAT 560
Qy 907 eLeuGluMetArgTyrSerGlyValGlyIleAspGluTyrTyrArgAsnGluInPheTr 927
Db 561 ATTGGAGCTTAGATGGAGCGGTGTGGCATTTGAAGATTGGTGGAGAAATGAGCAGTTTGT 620
```

```
QY 927 pValIleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValle 947
Db |||||
621 GGTATTGGTGACCTCTGCCCATCTTCGCAGTGTCCAGGTCTGCTGAAAGTGT 580
QY 947 uAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPh 967
Db |||||
681 GGCCTGGGATTGATPACCAACTTCACAGTTACCTCAAAGGCATCTGATGAGATGGCGACTT 740
QY 967 eAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIl 987
Db |||||
741 TGCTGAGCTATATGTGTTCAAGTGGACCAAGTTTGCTCATTCCTCCGACCACTGTTCTTGT 800
QY 987 eIleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSe 1007
Db :|||
801 CATTAACTGGTGGGAATGGTGGCAGGAATTCGTATGCCATTACAGTGGCTACCAATC 860
QY 1007 rTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrPr 1027
Db |||||
861 CTGGGGTCCGCTCTTTGGAAAGCTGTTCTCTCGATCTGGGTGATCCTCCTCACTCTCTACCC 920
QY 1027 oPheLeuLysGlyLeuMetGlyArgGlnAsn 1037
Db |||||
921 CCTTCTCAAGGTCATCGGAAGGCAAAAC 951
```

Search completed: August 22, 2004, 22:21:54
Job time : 6272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 13:37:22 ; Search time 935 Seconds
(without alignments)
4907.010 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGADALKSRHGAGDVQC.....VDPFTTRLAGNIQTGGINC 1080

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09900237/runat.18082004_081513_14627/app_query.fasta_1.1223
-DB=N_Geneseq_29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09900237.qcgn 1.1.649 @runat.18082004_081513_14627 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5434.5	94.1	3264	7 ADA69779	Ada69779 Rice gene
2	5423.5	93.9	3799	6 AAS16458	Aas16458 Corn cDNA
3	5421.5	93.8	3746	3 AAZ99512	Aaz99512 DNA encod
4	5421.5	93.8	3746	3 AAZ99527	Aaz99527 DNA encod
5	5421.5	93.8	3773	3 AAZ99494	Aaz99494 DNA encod
6	5169.5	89.5	3222	7 ADA69499	Ada69499 Rice gene
7	5143	89.0	3704	3 ADA99533	Aaz99533 DNA encod
8	4530	78.4	3198	3 AAC49550	Aac49550 Arabidops

ALIGNMENTS

RESULT 1

ADA69779

ID ADA69779 standard; DNA; 3264 BP.

XX AC ADA69779;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3102.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;

XX KW gene; ds.

XX OS Oryza sativa.

XX FN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

Aav06567 Arabidops
Aaz99500 DNA encod
Aaz99521 DNA encod
Aaz99506 DNA encod
Aaz99530 DNA encod
Aaz99515 DNA encod
Aaz99497 DNA encod
Aav06565 Arabidops
Aav06568 Arabidops
Aa67114 Pinus rad
Aaz99509 DNA encod
Aaz99524 DNA encod
Aaz99491 DNA encod
Aaz58263 Corn cell
Aaz58266 Corn cell
Aaz99518 DNA encod
Aaz99503 DNA encod
Aaz58268 Soybean c
Aav08373 Cellulose
Abz12754 Arabidops
Abz14725 Arabidops
Aav06566 Arabidops
Aac83798 Arabidops
Aas16455 Corn cDNA
Aav06563 Arabidops
Aac83247 Cellulose
Aav08372 Cellulose
Aav34432 Cotton ce
Aaa67145 Eucalyptu
Aac65448 Populus t
Aaz58270 Soybean c
Aav34433 Cotton ce
Aaz58271 Wheat cel
Aav08381 Cellulose
Aav06562 Arabidops
Aaz58269 Soybean c

9 4499 77.9 3614 2 AAV06567
10 3993.5 69.1 3725 3 AAZ99500
11 3993.5 69.1 3725 3 AAZ99521
12 3993.5 69.1 3725 3 AAZ99506
13 3981 68.9 3753 3 AAZ99530
14 3981 68.9 3753 3 AAZ99515
15 3981 68.9 3780 3 AAZ99497
16 3963.5 68.6 3603 2 AAV06565
17 3953.5 68.4 3673 3 AAV06568
18 3939 68.2 3851 3 Aa67114
19 3844.5 66.5 3786 3 AAZ58265
20 3844.5 66.5 3813 3 AAZ99509
21 3844.5 66.5 3813 3 AAZ99524
22 3820 66.1 3568 3 AAZ99491
23 3818 66.1 3776 3 AAZ58263
24 3818 66.1 3936 3 AAZ58266
25 3818 66.1 3969 3 AAZ99518
26 3818 66.1 3969 3 AAZ99503
27 3807.5 65.9 3517 3 AAZ58268
28 3777.5 65.4 3311 2 AAV08373
29 3666.5 63.5 3255 6 ABZ12754
30 3636 62.9 3255 6 ABZ14725
31 3633.5 62.9 3828 2 AAV06566
32 3583 62.0 3444 4 AAC83798
33 3535 61.2 2830 6 AAS16455
34 3518.5 60.9 8411 2 AAV06563
35 3514.5 60.8 7234 4 AAC83247
36 3482 60.3 3207 2 AAV08372
37 3477 60.2 3328 2 AAV34432
38 3461.5 59.9 3747 3 AAA67145
39 3337.5 57.8 3232 4 AAC65448
40 3234 56.0 2890 3 AAZ58270
41 2820 48.8 2306 2 AAV34433
42 2689 46.5 1734 3 AAZ58271
43 2674 46.3 2033 2 AAV08381
44 2629.5 45.5 2248 2 AAV06562
45 2627.5 45.5 2125 3 AAZ58269

PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

XX Claim 6; SEQ ID NO 3102; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 3264 BP; 784 A; 751 C; 852 G; 877 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3264
 Score: 5434.50 Matches: 998
 Percent Similarity: 96.60% Conservative: 52
 Best Local Similarity: 91.81% Mismatches: 30
 Query Match: 94.06% Indels: 7
 DB: Gaps: 2

US-09-900-237A-30 (1-1080) x ADA69779 (1-3264)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 Db 1 ATGGACGGCGACGCGGATGCGGTGAAGTCGGGAGGCGACGGGAGCGCGGCGTCCAG 60
 QY 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 Db 61 ATCTGGCGGACGGCGTGGCGACGACGGCGGGCGGCGGTGTTCGCCGCTCGCACGTC 120
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 Db 121 TGGCGCTTCGGGTGTGGCGCCCTGCTACGAGTACGAGCGCAAGGATGGCACCAGGCT 180
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 Db 181 TGCCCCCAGTGCAGACCAAGTACAAGCGCCACAAGGGAGCGCGCGATCCCGTGGGAG 240
 QY 81 GluGlyAspAspThrAspAlaAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 Db 241 GAAGCGGAGGATGATGCTGTGATGATGTCAGTGACTACAACTACCTCGCATCTGGCAGT 300
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
 Db 301 GCCGACCCAGAGCAGAGGATGCTGATAGGATCGGCAGTTGGCGCATGAATGCTGGGGT 360
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 Db 361 GGTGGAGACCTCGCGCGTCCCAAGTATGACAGTGGCGAGATCGGGCTCCACCAAGTATGAC 420
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 Db 421 AGTGGCGAGATCCCTCGGGGATACATCCCTTCAGTACTAATAGCCAGATCTCGGGAGAA 480
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 Db 481 ATCCCTGGAGCTTCCCTGATCATCATATGATGTCTCTACCGGAAACATTGGCAAGCGT 540
 QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
 Db 541 GCTCCATTCCTATGTGAACCATTCACCAATCCATCAAGGGAGTTCCTCGCAGCATT 600
 QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyVala 220
 Db 601 GGAAATGTTCCCTGGAAAGAAGAGTTGATGGCTGGAAACTGAGACAGGCAAGGGAGCA 660
 QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGlyArgAlaIaThrAspIle 240

661 ATTCCCATGACCAACGGGACAAAGCATTTGCCCTTCTGAAGGTCGGGGAGTTGGTGATATC 720
 QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 Db 721 GATGATCCACTGATTAACAATATGAAGATGCTTACTGAAATGATGAATACTGCCAGCTT 780
 QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 Db 781 CTCTCTAGAAAAGTTCCCTTCTTCAATCCAGAAATAATCCCTACAGAAATGTCATTGTT 840
 QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 Db 841 CTGCGTTGGTTGTTCTAAGCATTTTCTTCACTACCGTATTTACGAATCCTTGGCAAT 900
 QY 301 AlaTyrProLeuTrpLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
 Db 901 GCGTATCCGCTCTGGCTTTTATCTGTATATGTGAGATTGGTTGCTTGTCTCGGATA 960
 QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrIleuAspArgLeuAla 340
 Db 961 TTGGATCAGTTCCTCCGAAGTGGTTTCCAAATCAACCGTGAAACCTACCTTGATAGGCTGGCA 1020
 QY 341 LeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheValSer 360
 Db 1021 TTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTTTGTGAGT 1080
 QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
 Db 1081 ACAGTCGACCCCATGAAGAACTCTCTTGTGTACTGCGCAATACCGTGTGTCCTTCCT 1140
 QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 Db 1141 GCTGTTGATTACCCAGTCAGCAAGGTCTCTGCTATGTATCTGACGATGGTGTGCAATG 1200
 QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 Db 1201 CTGACTTTTGATGCAATTGGCTGAGACTTCAGAGTTTGTAGAAAGTGGGTTCCTTCGTT 1260
 QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 Db 1261 AAGAATATACCAATTGAGCCAGAGCTCTCTGAGTGTACTTCTCCCAGAAAATCGATTAC 1320
 QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
 Db 1321 TTGAAGACAAAGTCCACCTTCATTTGTTAAAGACCGTCGTGCCATGAAGAGAGATAT 1380
 QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 Db 1381 GAAGAAATCAAAGTTAGGATAAATGGCTTGTGTGCTAAGGACACAGAAAGTCCCGGAGGAG 1440
 QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
 Db 1441 GGATGATCATGCAAGATGGCACACCATGCCAGGAAACAATACTAGGGACCATCTCTGGA 1500
 QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
 Db 1501 ATCATTTGAGTTTCTTCTGTCACAGTGTGGCTTGTATACCGAGGCTATGAGCTTCCC 1560
 QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
 Db 1561 CGCTCGTCTACGTATCTCGTGAAGACGCTCTGGGTTCAGCACCAAGAAAGCTGGT 1620
 QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
 Db 1621 GCCATGAATGCTCTTGTTCGTGCTCAGCTGTGCTTACCAATGGACAGTACATGTTCAAT 1680
 QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 Db 1681 CTTGATTGTTGATCACTACATCAACCAACAGCAAGGCTCTCGCGGAAGTATGTGCTTCTT 1740
 QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600

```

Db 1741 ATGGATCCAAACCTAGGAAGAGTGTCTGTATGTTTCAGTTCCTCCACAAAGGTTTCGATGTT 1800
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPheAspPheAsnLeuArg 620
Db 1801 ATTGATAGGATGATGATATGCGACAGAACACTGTGTGTTTTCATATTAACATTGAGG 1860
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 1861 GGTCTGTATGTCATCCAAAGACCAAGTTTATGTGGGAACCTGTGTGTATTCAACAGAACT 1920
QY 641 AlaIleTyrGlyTyrGluProProlleLysAlaLysLysProGly---PheLeuAlaSer 659
Db 1921 GCTCTATATGTTATGAACCCCAATTAAGCAGACAGAAAGGAAGTTCTTCTGTCATCA 1980
QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 1981 CTATGTGGGGCGGGAAGGACAGCAAGTCAAAGAAAGAGTTCGGCAGGAAGAG 2040
QY 680 SerAsnLysHisValAspSerValProValPheAsnLeuGluAspIleGluGly 699
Db 2041 TCAAAACAAGCAGCTGGACAGTCTGTGCCAGTTTTCATCTTGAAGATATAGAGGAGGT 2100
QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2101 GTTGAAGGTGCTGGATTCGATGATGAGAAATCACTTCTTATGCTCAAATGAGCTTGGAG 2160
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
Db 2161 AAGAGATTGGCCAGTCTGACGCGTTGTGCTCCACTCTCATGGAATATGTTGGTGT 2220
QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2221 CCTCAATCTGCAACCCCAAGATCTCTTTGAAAGAGCTATCCATGATGATAGTTGTGGC 2280
QY 760 TyrGluAspLysSerSerGluTyrGlyThrGlu-----IleGlyTyrIle 773
Db 2281 TATGAGGACAAGACGAATGGGGAACCTGAGGCTATGCAACTGCAATGATTGGGTGGATC 2340
QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
Db 2341 TACGGTTCGGTGACAGAGATATCTCACTGATTCGAAGATGCAATGCGCGAGGTGGAGA 2400
QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
Db 2401 TCAATCTACTGATGCCCAAGGCCACCTTTCAGGGGCTGCTGCTCTATCAATCTTCA 2460
QY 814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833
Db 2461 GATCGTCTTAACCAAGTGCITTCGGTGGCACTTGGTCTGTGAAATCTTTTTCAGTGC 2520
QY 834 HisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
Db 2521 CATTTGCCATATGATCGGCTATGGAGGACGCTTAAGTTCTTGGAGAGATTTGCCATC 2580
QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuLeuValTyrCysIleLeuPro 873
Db 2581 ATCAACACCACATATTTATTCATTGACATCGATCCCGCTTCTCATATACTGTGTTTGCCT 2640
QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
Db 2641 GCTATCTGTTTGTCTCACTGGGAAGTTCATCATCCAGAGATTAGCACTTTGCTAGTATT 2700
QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913
Db 2701 TGGTTCACTCTCTCTTCATTTCAATTTTGGCCACTGGTATCTCTTGAGATGAGTGGAGT 2760
QY 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
Db 2761 GGTGTTGGCATCGATGATGAGTGGTGGAGAAATGAACAGTTCTGGGTTATGGAGGTATATCT 2820
QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
Db 2821 GCGCATCTTTTGGCGTCTTCCAGGGTCTCTCTCAAGGTGCTTGTGATCGACACCAAT 2880
QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
Db 2881 TTCACTGTCACTCAAGGCTTCTGATGAAGATGGCACTTTTGTCTGAGCTCTACATGTTTC 2940
QY 974 LysTyrThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal 993
Db 2941 AAGTGACAAACGCTTCTCATCCACGACGACATCTTGATCATTAACTGTGCTGTGTT 3000
QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly 1013
Db 3001 GTTGTGTTATCTCATCGGATCAACAGCGCTACCATTCATGGGACCGCTCTTTGCG 3060
QY 1014 LysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
Db 3061 AAGCTCTCTTTTGGCTTCTGGGTGATTGTCCACTTGTACCCCTTCTCAAGGGTCTTATG 3120
QY 1034 GlyArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIle 1053
Db 3121 GGTGCGCAAAACCGCACCTCCGACCATCGTTGTGTTGGGCAATCTTCTGCTTTCGATC 3180
QY 1054 PheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
Db 3181 TTCTCATTTGCTGTGGTTTCGATCGATCCATTCACACCCGTGTCCACCGGCCCATACC 3240
QY 1074 GlnThrCysGlyIleAsnCys 1080
Db 3241 CAAACATGGCATCAACTGC 3261
RESULT 2
AAS16458
ID AAS16458 standard; cDNA; 3799 BP.
XX
AC AAS16458;
XX
DT 14-FEB-2002 (first entry)
XX
DE Corn cDNA encoding cellulose synthase Cqrae19/cesa-19.
XX
KW Corn; ss; cellulose synthase; Cdpqs45; cesa-3; Cqrae19; cesa-9;
KW stalk quality; improved stand; silage; pericarp; kernel hardening;
KW handling ability; transgenic plant.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 238..3799
FT /*tag= a
FT /product= "Cellulose synthase"
XX
XX WO200179516-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US011951.
XX
XX 14-APR-2000; 2000US-00550483.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG;
XX
XX WPI; 2002-041338/05.
XX
XX P-PSDB; AAU10496.
XX
XX New cellulose synthase polypeptides and polynucleotides, useful in
XX improving stalk quality or silage, and in increasing concentration of
XX cellulose in the pericarp, hardening the kernel for improved handling
XX ability.
XX
XX Claim 4; Page 80-85; 88pp; English.
XX
XX The invention relates to isolated nucleic acids encoding two cellulose
CC

```

CC synthase proteins from corn, Cdp945 (cesa-3) and Cqrae19 (cesa-9). Also
 CC disclosed are a recombinant expression cassette comprising the
 CC polynucleotide (operably linked to a promoter) a host cell comprising the
 CC recombinant expression cassette and a transgenic plant comprising the
 CC recombinant expression cassette. The nucleic acid is useful in the
 CC improvement of stalk quality for improved stand or silage, and in the
 CC increased concentration of cellulose in the pericarp, hardening the
 CC kernel, and thus improving its handling ability. The nucleic acids may
 CC also be used as probes or amplification primers in the detection,
 CC quantification or isolation of gene transcripts, as probes in detecting
 CC deficiencies in the level of mRNA, for detecting gene mutations or
 CC allelic variants, for monitoring up regulation of expression or changes
 CC in enzyme activity in screening assays, for site directed mutagenesis,
 CC and in sense or antisense suppression of one or more genes in a host
 CC cell, tissue or plant. The polypeptides may be used in assays for enzyme
 CC agonists or antagonists, and as immunogen or antigen to obtain antibodies
 CC specifically immunoreactive with the protein. The present sequence
 CC encodes a corn cellulose synthase of the invention
 XX

SQ Sequence 3799 BP; 925 A; 931 C; 995 G; 944 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 0 Length: 3799
 Score: 5423.50 Matches: 996
 Percent Similarity: 96.48% Conservative: 46
 Best Local Similarity: 92.22% Mismatches: 37
 Query Match: 93.86% Indels: 1
 DB: 6 Gaps: 1

US-09-900-237A-30 (1-1080) x AAS16458 (1-3799)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 DB 238 ATGAGGGGACCGGACCGCGTGAAGTCGGGAGGCGCGGGAGGCGAGGTGTGCCAG 297
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 298 ATCTGGCGGATGGCGTGGGCACCTACCGCGGAGGAGACGCTCTTCACCGCTCGCAGTCC 357
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 DB 358 TGGCGGTTCCGGTGTGCGCCCTGCTACGAGTACGAGCGCAAGCGGCAAGCAGCG 417
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 418 TGCCCCCAGTGCAGAAACAAAGTCAAGCGCCCAAGGGGAGTCCAGCGATCCCGAGGGAG 477
 QY 81 GluGlyAspThrAspAlaAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 DB 478 GAAGGACGACATCTGATGCCGATGATGCTAGCGACTTCAACTACCTCGCATCTGGCAAT 537
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
 DB 538 GACGACCAAGCAGAGAAATGTCTGACAGGATGCGCAGCTGGCGCATGAATGTGGGGGC 597
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 DB 598 AGCGGGATGTGGCGCCGCCCAAGTATGACAGTGGTGAGATCGGGCTTACCAGTACGAC 657
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 DB 658 AGTGGTGAGATCCCTCGGGGATACATCCCGTCAGTCACTAACAGCCAGATTCGGGAGAA 717
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 DB 718 ATCCCTGTGTCTCCCTCGACCATCATATGATGCTCTACTGGGAAATTCGGCAGGGCGC 777
 QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
 DB 778 GCCCATTTTCCCTATATGAATCATCAATCCGTGAGGGAAATCTCTGGTAGCGT 837
 QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220

DB 838 GCGAATGTTGCTGGAAGAGAGAGGGTGTGATGCTGGAAATGAACGACGACAGGGAACA 897
 QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
 DB 898 ATTCCATGACGAATGGCACAGCATGTCTCCTCTGAGGGCGGGGTGTGGTGATAT 957
 QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 DB 958 GATGCATCAACTGATTACCAACATGGAAGATGCTTATTAAACGATGAAATCGCCAGCCT 1017
 QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 DB 1018 CTATCTAGAAAGTTCCACTTCTCTCCAGATTAATCCATACAGATGTGTCTATGTG 1077
 QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 DB 1078 CTACGATTGATTGTTCTAAGCATCTTCTTGCACTACCGGATCACAAATCTGTGCGTAAT 1137
 QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
 DB 1138 GCATACCCACTGTGGCTTCTATCTGTATATGTGAGATCTGTGTTCTCTTCTCTGATA 1197
 QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
 DB 1198 TTGGATCAGTTTCCAAAGTGGTTTCCATCAACCGCGAGACTTACCTGTATACACTGCA 1257
 QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
 DB 1258 TTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTGTGCTGTGTGACATTTTGTGAGT 1317
 QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
 DB 1318 ACTGTGCGACCAATGAAGGAGCCTCTCTGTCTACCTGCCAATACCGTGCTATCCATTC 1377
 QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 DB 1378 GCTGTGACATCTCTGTGGATAAGGTCTCTTGTAATGATCTGATGATGAGCTGCTATG 1437
 QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 DB 1438 CTGACATTTGATGCATAGCTAGACTTCAGAGTTTGTCTAGAAATGGGTGCCATTTGTT 1497
 QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 DB 1498 AAGAAGTACAACTGAACCTAGAGCTCTCTGAATGGTACTTCTCCCAAGAAATGATATAC 1557
 QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
 DB 1558 TTGAAGGACAAAGTGCACTTCTTCTTTGTTAAAGACCGCGGGCCATGAAGAGAAATAT 1617
 QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 DB 1618 GAAGAAATCAAAATTAGGGTAAATGGCTTGTGCTAAGGCACAAAAAGTCCCTGAGGAA 1677
 QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
 DB 1678 GGATGATCATGCAAGATGGCACCATGGCCAGGAAACAATACAGGACCATCTCTGGA 1737
 QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
 DB 1738 ATGATTCAGGTTTTCTTGTGTGTCAGTGTGTCTTGTACTGAGGGTAAATGAGTACCC 1797
 QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
 DB 1798 CGTTTGGTCTATGTTTCTCGTGAAAAACGTCCTGGATTCAGCATCAACAAGAAAGCTGGT 1857
 QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
 DB 1858 GCCATGAATGTCTGTCCGGCTCTCAGCTGTGCTTCAATGGAACAATACATGTGTGAAT 1917
 QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 DB 1918 CTTGATGTGATCACTACATCAACACAGTAGGCTCTCAGGGAAGCTATGTGCTCTCTT 1977

Db 1995 CTTGATTGTGATCACTAATTAACAACAGTAGGCTCTCAGGAGACTATGTCTTCCT 2054
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db 2055 ATGACCCCTAACCTAGGAGAGTGTCTGCTACGTCCAGTTTCCCCAGAGATTGATGCG 2114
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db 2115 ATTGACAGGATGATCGATATGCCACAGGACACCGTGTTTTCGATATTAACTTGAGA 2174
QY 621 GlyLeuAspGlyIleGlnClyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 2175 GGTCTTGATGTCATCCAGGACCAAGTATATGCGAACTGCTGTGTTTCAACCGACA 2234
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2235 GCTCTATATGTTTATGACCCCAATTAAGCAGAGAGGGTGTGTTTCTGTCTCACTA 2294
QY 661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer 680
Db 2295 TGTGCGGTAGAGAGAGGCAAGCAATCAAGAGAG---GGCTCGACACAGAGATGCG 2351
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2352 CAGAAGCATGTGACAGTCTGTGCCAGTATTCAACTTGAAGATATAGAGAGGAGT 2411
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2412 GAAGGCGCTGGAATTTGACGACGAGAAATCACTCTTATGTCTCAATGAGCTGAGAG 2471
QY 721 ArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2472 AGATTGCGCAGTCCGACAGCTTTGTGCTTCCACTCTGATGGAGTATGGTGTCT 2531
QY 741 GlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2532 CAGTCCGCACTCCGGAGTCTCTCTGAAGAAGCTATCCATGTTATAAGCTGTGGCTAT 2591
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAsp 780
Db 2592 GAGGACAAAGACTGAATCGGAACCTGAGATCGGTGGATCTACGGTCTCTGACAGAGAC 2651
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLys 800
Db 2652 ATTCTACCGGATTCAGATGCACGCGAGGCTGGCGTGCATCTACTGATGCCCAAG 2711
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2712 CGGCCAGCTTCAAGGGGTCTGCCCCCATCAATCTTTCGAGACCGCTCTGAACCAAGTCTC 2771
QY 821 ArgTrpAlaLeuGlySerValGluIleuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2772 CGGTGGGCTCTTGGGTCCGTGGAGATCCTCTTCAGCCGGCACTGCCCTCTGTGTACGCG 2831
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2832 TAGGGGGCGGCTCAGTTCTTGGAGAGATTCGGTACATCAACACCACTATCCCG 2891
QY 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2892 CTCAGTCCATCCGCTTCTCATCTACTGATCTCTGCCCGCATCTGTCTGCTCACCGGA 2951
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db 2952 AAGTTTCATATCCAGAGATCAGCAACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCA 3011
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTyr 920
Db 3012 TCGATCTTCGCCACGGCATCTTGGAGATGAGTGGAGCGGGTGGCGATCGAGAGTGG 3071
QY 921 TrpArgAsnGluGlnPheTrpValIleGlyLysSerAlaHisLeuPheAlaValPhe 940

Db 3072 TGGAGRAAGCAGCAGTCTCTGGGTGATCGGGGCATCTCCGGCACCTCTTCGCCGTGTT 3131
QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerIysAla 960
Db 3132 CAGGCGCTGTCTCAAGGTCTCGCGGCATCGACACCAACTTCACGTCACCTCCAAGGCC 3191
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db 3192 TCGACGAGGACGCGACTTCGCGAGCTGTACATGTTCAAGTGGACAGCCTCTCTGATC 3251
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3252 CGGCCACCACTCTGATCATCACTGTGCGGCTGCTCGCGGCATCTCTTACGCC 3311
QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db 3312 ATCAACAGCGGATACCACTGTGCGGCGCGCTCTTCGSCAAGCTCTTCTCGCTCTCG 3371
QY 1021 ValIleValHistuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3372 GTCATGTCACCTGTATCCCGTCTCTCAAGGCGCTCATGGGCGAGCAGAACCGCACCCG 3431
QY 1041 ThrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3432 ACCATGCTGCTGCTGCGCATCTGCTGCGTCCATCTTCTCTGCTGTGGTTCGC 3491
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3492 ATCGACCCCTTCAACACCGCGTCACTGGCGCGGATACCCAGACGTGTGGCATCAACTGC 3551
RESULT 4
AAZ99527 standard; DNA; 3746 BP.
XX
AC AAZ99527;
XX
DT 03-JUL-2000 (first entry)
XX
DB DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 321..3449
FT /tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1800..1802, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified amino acid"
XX
FN W0200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
DR P-PSDB; AAY84119.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
PS Claim 1; Page 176-181; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 XX
 SQ Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 0 Length: 3746
 Score: 5421.50 Matches: 998
 Percent Similarity: 96.67% Conservative: 46
 Best Local Similarity: 92.41% Mismatches: 33
 Query Match: 93.83% Indels: 3
 DB: 3 Gaps: 2

US-09-900-237A-30 (1-1080) x AAZ99527 (1-3746)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 DB 321 ATGGAGGGGACGCGGCGTGAAGTCGGGAGGCGGGTGGCGACAGGTGTGCCAG 380
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 381 ATCTGGGGGACGGCGTGGGACACCGCGGAGGGGAGCGTCTTCGCCCGCTCGGACGTC 440
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlnAla 60
 DB 441 TCGGGGTTTCGGGTGTGGCGCCCTCTACGAGTACGAGCGGCAAGACGGCGACGCGAGCG 500
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 501 TCGCCCCAGTGCACAGCAGTACAGCGCCACAGGGGAGCGCGCGATCCGTGGGGAG 560
 QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 DB 561 GAAGGAGACGACACTGATGCCGAT-----AGCGACTTCAATTACCTTCGATCTGGCAAT 614
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
 DB 615 GAGGACCAAGACAGAGATTTGCCGACAGATTCGGCAGCTGGCCCATGACGTTGGGGGC 674
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 DB 675 AGCGGGGATGTTGTCGCCCAAGTATGACAGTGGCGGAGATCGGGCTTACCAAGTATGAC 734
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 DB 735 AGTGGCGGAGATTCCTCGGGGATACATCCCATCATGTCCTAACACGACGAGATCTCAGGAGAA 794
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 DB 795 ATCCCTGGTGCTTCCCTCGACCATCATATGATGTCCCAACTGGGACATTCGCAAGCGT 854
 QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200

DB 855 GCTCATTTTCTCTATGTGAACCATTCGCCAAATCCGTCAAGGGAGCTTCTCTGGTAGCATT 914
 QY 201 GlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
 DB 915 GGGAAATGTCTGAAAGAGAGGGTTGATGCTGGAAAAATGAACACGACGACNAGGGGAGC 974
 QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
 DB 975 ATTCCTCATGACGAATGGCACAAGCATTCCTCCCTGAGGGTCCGGGTGTTGGTGATATT 1034
 QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 DB 1035 GATGCATCAACTGATTACAAACATGCAAGATGCTTTATTGAACACGACGAACCTGCAGCGCT 1094
 QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 DB 1095 CTATCTAGAAGATTCACCTTCCTCCAGGAATAATCCATACAGATGTGTCATTGG 1154
 QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 DB 1155 CTGCGATTGATTGTTCTAAGCATCTCTTGCACTACCGTATCAAAATCCTGTGGGCAAT 1214
 QY 301 AlaTyrProLeuTyrLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle 320
 DB 1215 GCATACCCCATTTATGGCTTCTATCTGTTATATGTGAGATCTGGTTTGTCTCTTTCGTGGATA 1274
 QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
 DB 1275 TTGGATCAGTTCCTTAAGTGGTTTCAATCAACCGGGAGACGTACTTGTATAGGCTGGCA 1334
 QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
 DB 1335 TTAAGGTATGACCGGGAAGTGCAGCATCTCAGTTGGCTGCTGTGACATTTTTCGTCAST 1394
 QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
 DB 1395 ACAGTCGACCCCAATGAAGAGACCTCTCTGTGCACGCAATACCGTGTCTATCCATCTT 1454
 QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 DB 1455 GUTGTGGATTACCTGTGGTAAAGTCTCTGCTATGATATCTGTATGATGATGAGCTGGAG 1514
 QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 DB 1515 CTGACATTTGATGACTAGCTCAGACTTCAGAGTTTGTAGAAAAATGGGTACCATTTGTT 1574
 QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 DB 1575 AAGAAGTACAACTGAACCTAGAGCTCTGAAATGGTACTTCTCCAGAAAATTGATTAC 1634
 QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaAlaMetLysArgGluTyr 460
 DB 1635 TTGAGGACAAAGTCACCTTCATTGTTAAAGCCGCGGGCCATGAAGAGAAATAT 1694
 QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 DB 1695 GAAGAATTCAAAGTTAGGCTAAATGGCTTGTGCTTAAGCACAGAAAGTTCCTGAGGAA 1754
 QY 481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
 DB 1755 GGATGATCATGCAAGATGGCACACCATGCCAGGAAACAATACCCGGGACCACTCTGGA 1814
 QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
 DB 1815 ATGATTGAGTTTCTTGTGTCAGTGTGGCTTGATATCGAGGGCAATGAGCTACCC 1874
 QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIleLysLysAlaGly 540
 DB 1875 CGTTTGGTCTAATGTTCTCGTGAAGAGCGTCTCTGGATTCCAGCATCACAAAGAGCTGT 1934
 QY 541 AlaMetAsnAlaLeuValValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
 DB 1935 GCCATGAATGCTCTTGTGCTGCTGCTTACCAATGACCAATACATGTTGAAT 1994

QY 561 LeuAspCysAspHisTyrIleAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 Db 1995 CTTGATTGTGATCACTACATTAAACAGATAGAGCTCTCAGGAAAGCTATGCTTCCTT 2054
 QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
 Db 2055 ATGGACCCCTAACTAGGAAGAGTGTCTGCTACGTCCAGTTTCCCGCAGAGATTGATGCG 2114
 QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
 Db 2115 ATTGCAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTTCGATATTAACTTGAGA 2174
 QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
 Db 2175 GGTCTTGATGGATCCAGGACAGTTTATGTCCGAACCTGGCTGTGTTTCAACCGAACA 2234
 QY 641 AlaIleTyrGlyTyrGlnProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
 Db 2235 GCTCTATATGTTATGAGCCGCCCAATTAAAGCAGAAGAGGTGGTTTCTTGTCACTACTA 2294
 QY 661 CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer 680
 Db 2295 TGTGCGGTGAGGAAGAGCAAGCAAGCAAAATCAAAGAAG---GGCTCGGCAAGAAGAAGTCG 2351
 QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGlnGlyVal 700
 Db 2352 CAGAAGCATGTGGACAGTTCTGTGCGAGTATTCAACCTTGAGATATAGAGGAGGAGTT 2411
 QY 701 GluGlyAlaGlyPheAspAspGlyLysSerValLeuMetSerGlnMetSerLeuGluLys 720
 Db 2412 GAAGGCGCTGATTTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCTGGAGAAG 2471
 QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
 Db 2472 AGATTGTGCCAGTCGCGAGCGTTGTTCCTCCACTCTGATGGAGTATGGTGGTTCCT 2531
 QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
 Db 2532 CAGTCCGCACTCCGGAGTCTCTCTGAAAGAGCTATCCATGTTATAGCTGTGGCTAT 2591
 QY 761 GluAspLysSerGluThrGlyLeuIleGlyTyrIleTyrGlySerValThrGluAsp 780
 Db 2592 GAGGACAAGACTGAATGGGAACCTGAGATCGGGTGGATCTACGGTTCGTGACAGAGAC 2651
 QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
 Db 2652 ATTCTACCGGATTCAAGATGACGCGAGGCTGGCGGTGCGATCTACTGATGCCCAAG 2711
 QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
 Db 2712 CGGCCAGCTTCAAGGGGTCTGCCCCCACTCAATCTTTCGGACCGTCTGACAGGTGCTC 2771
 QY 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTrpGly 840
 Db 2772 CGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCGGCACTGCCCTCTGTGTACGGC 2831
 QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
 Db 2832 TACGGAGGGCGGCTCAAGATTCTCTGGAGAGATTCGCGTACATCAACACCCATCTACCCG 2891
 QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
 Db 2892 CTCAGTCCATCCCGCTCTCTACTACTGATCTTCCCGGCACTGTCTGCTCACCGGA 2951
 QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
 Db 2952 AAGTTCATCTCCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3011
 QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
 Db 3012 TCGATCTTCGCGACGGCATCTCGAGATGAGGTGGAGCGGGTGGCGATCGAGAGTGG 3071

QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisIleuPheAlaValPhe 940
 Db 3072 TGGAGGAACGAGCAGTTCTGGGTGATCGGGGCACTTCGCGCACCTCTTCGCGGTGTTCC 3131
 QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
 Db 3132 CAGGGCTGTCTCAAGGTGTGCGCGGCATCGACCACTTCACCGTCACTCCCAAGGCC 3191
 QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
 Db 3192 TCGGACGAGGACGGGACTTCGCGGAGCTGTACATGTTCAAGTGGACGAGCTCCTGATC 3251
 QY 981 ProProThrThrIleIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
 Db 3252 CGCCCAACACCATCTGATCATCAACCTGGTGGCGCTGTCGCGGCATCTCTCTACGCC 3311
 QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
 Db 3312 ATCAACAGGGATACCACTGCTGGGCGCGCTCTTCGGCAAGCTCTTCTCGCCTTCTGG 3371
 QY 1021 VallIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
 Db 3372 GTCATGCTCCACTGTACCGTTCCTCAAGGCGCTCATGGGCGAGCAGAACCGCACCCG 3431
 QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
 Db 3432 ACCATGCTGCTGCTGCTGGCCATCTGCTGGCGTCCATCTTCTGCTGTGGGTTCGC 3491
 QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
 Db 3492 ATCGACCCCTTACCACCGCGTCACTGGCGCGGATACCCAGACGTGTGGCATCAACTGC 3551

RESULT 5

AAZ99494

ID AAZ99494 standard; DNA; 3773 BP.

XX AAZ99494;

XX 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

FH Key Location/Qualifiers
 CDS 337..3565FT /*tag= a
 FT /product= "cellulose synthase"

FT /transl_except= (pos: 1817..1819, aa: Xaa)

FT /note= "no termination codon given; Xaa is an unspecified
 FT amino acid"

XX WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US018760.

XX 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPT; 2000-224343/19.

XX P-PSDB; AAY84108.

PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.

XX
PS
XX
Claim 1; Page 90-94; 119pp; English.
The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting or mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein

XX
SQ Sequence 3773 BP; 863 A; 988 C; 1017 G; 904 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 0 Length: 3773
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 3 Gaps: 2

US-09-900-237A-30 (1-1080) x AAZ99494 (1-3773)

XX	181	AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle	200
PS	872		931
XX			
QY			
DB			
QY	201	GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyVala	220
PS	932		991
XX			
QY	221	IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle	240
PS	992		1051
XX			
QY	241	AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro	260
PS	1052		1111
XX			
QY	261	LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal	280
PS	1112		1171
XX			
QY	281	LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn	300
PS	1172		1231
XX			
QY	301	AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle	320
PS	1232		1291
XX			
QY	321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
PS	1292		1351
XX			
QY	341	LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer	360
PS	1352		1411
XX			
QY	361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380
PS	1412		1471
XX			
QY	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSerMet	400
PS	1472		1531
XX			
QY	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
PS	1532		1591
XX			
QY	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
PS	1592		1651
XX			
QY	441	LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr	460
PS	1652		1711
XX			
QY	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
PS	1712		1771
XX			
QY	481	GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	500
PS	1772		1831
XX			
QY	501	MetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeuPro	520
PS	1832		1891
XX			
QY	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly	540
PS	1892		1951
XX			
QY	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560

DB 1952 GCATGAATGCTCTTGTCTGTCTCAGCTGTCTTACCAATGGAACAATACATGTTGAAT 2011
QY 561 LeuAspCysAspHisfyrilleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
DB 2012 CTTGATTTGTGATCAGTACATACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGCTTCCTT 2071
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
DB 2072 ATGGACCTTAACTAGGAAGGAGTGTCTGTCTAGCTCCAGTTTCCCAGAGATTCGATGGC 2131
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
DB 2132 ATTGACAGGAATGATCATGATATCAACAGCAACACCGCTGTTTTCGATATTAACTTGACA 2191
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
DB 2192 GGTCTTGATGGATCCAGGACAGTTTATGTCGGAACCTGGCTGTGTTTCAACCGAACA 2251
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
DB 2252 GCTCTATATGTTATGAGCCCCCAATTAAAGCAGAGAAGGGTGGTTCTTGTATCATCTA 2311
QY 661 CysGlyGlyLysLysAlaSerLysSerLysArgSerSerAspLysLysSer 680
DB 2312 TGTGGCGGTAGGAAGCAAGCAAAATCAAGAAG---GGCTCGACAAAGAAGTGC 2368
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
DB 2369 CAGAAGCATGTGGAAGTCTGTGTCAGTATTCAACCTTGAAGATATAGAGGAGGAGTT 2428
QY 701 GluGlyAlaGlyPheAspAspGlyLysSerValLeuMetSerGlnMetSerLeuGluLys 720
DB 2429 GAAGCGCTGATTTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCTGGAGAG 2488
QY 721 ArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
DB 2489 AGATTGGCCAGTCGCGAGCGTTTGTGCTCCACTCTGTAGTGGATGTTGGTGTCTCT 2548
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
DB 2549 CAGTCGCAACTCCGGAGTCTCTCTGAAGAAGATATCCATGTTATAGCTGTGGCTAT 2608
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrPheLysSerValThrGluAsp 780
DB 2609 GAGGACAGACTGATGGGGAAGTGAATCGGTGGATCTACGGTCTGTGACAGAAAC 2668
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
DB 2669 ATTCTACCGGATTCAGATGCGCGAGGCTGGCGGTGATCTACTGATGCCCAAG 2728
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
DB 2729 CGGCCAGCTTTCAAGGGGTGTGCCCCCATCAATCTTTCGAGCCGCTCTGAACAGGTGCTC 2788
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
DB 2789 CGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCCGGCACTGCCCTGTGGTACGGC 2848
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrIleTyrPro 860
DB 2849 TACGAGGGCGGCTCAAGTCTCTGGAGAGATTCGGTATACCAACACCATCTACCCG 2908
QY 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
DB 2909 CTCAGTCCATCCCGCTTCTCATCTACTGCACTCTGCGCGCATCTGTCTCACCGGA 2968
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
DB 2969 AAGTTTCATCTCCAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3028
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920

DB 3029 TCGATCTTCGCCACGGGCATCTCTGGAGATGAGGTGAGCGGGGTGGCATCGACGAGTGG 3088
QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
DB 3089 TGGAGGAACGAGCAGTCTTGGGTGATCGGGGCACTCTCCGCGCACCTCTTGGCCGTTTC 3148
QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
DB 3149 CAGGGGCTCTCTCAAGGTGCTGGCCGCAATCGACACCACTTCACGTCACCTCCAGGGC 3208
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
DB 3209 TCGGACGAGGACGGGCACTTTCGGGAGCTGTACATGTTCAAGTGGACGACGCTCTGTATC 3268
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
DB 3269 CGGCCACCACTCTGATCATCAACCTGCTGGCGCTGCTCGCGGCATCTCTTCGCTTCTGG 3328
QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
DB 3329 ATCAACAGCGGATACAGTCTGGGGCCCGCTCTTCGGCAAGCTCTTCTTCGCTTCTGG 3388
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
DB 3389 GTCATCTGCTCCACTGTATCCCGTTCCTCAAGGGCTCTCAAGGGCCATCTTCTTCTGTTGGTTCGC 3448
QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
DB 3449 ACCATCGTCTGCTCTGGCCATCTCTGCTGGCGTCCATCTTCTTCTGTTGGTTCGC 3508
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
DB 3509 ATCGACCCCTTTCACACCCGCTCTCACTGGCCCGGATACCCAGACGTGTGGCATCACTGC 3568
RESULT 6
ADA69499
ID ADA69499 standard; DNA; 3222 BP.
XX ADA69499;
AC ADA69499;
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 2822.
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB0011105.
XX
PR 22-JUN-2001; 2001WO-IB0011105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 2822; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX
 SQ Sequence 3222 BP; 805 A; 711 C; 805 G; 901 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3222
Score:	5169.50	Matches:	950
Percent Similarity:	94.52%	Conservatives:	67
Best Local Similarity:	88.29%	Mismatches:	54
Query Match:	89.47%	Indels:	5
DB:	7	Gaps:	4

US-09-900-237A-30 (1-1080) x ADA69499 (1-3222)

QY	6	AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGly	25
DB	4	GACGGTGCAGAAATCGGGGAAGCAA-----TGCCATGTCTGCCAGATCTCGCGCGACGGC	57
QY	26	LeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysargPheProVal	45
DB	58	GTGGGACGCGCGCGGACGCGAGCTCTTCACCGCTCGCAGCTCTGGGGTCTCCGGTG	117
QY	46	CysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLys	65
DB	118	TGCGCGCCATGCTACGAGTACGAGCGCAGGACGCGCAGCGCGCGCGCGAGTGAAG	177
QY	66	ThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlyAspAspThr	85
DB	178	ACCAAGTACAAGCGCACAAAGGGGAGCGCCCGCATATCTGGGGATGAAGCGATGATGT	237
QY	86	AspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspGlnLysGln	105
DB	238	GATGCGGATGATGCTAGTGTGTAATATCCAACTCCCGCAACCGAGACCATATAGCAC	297
QY	106	LysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySerGlyAsnValGly	125
DB	298	AAAGTTGCCAGAGGATGCTCACCCTGGCGCATGACTCTGGAGGAATGATATTGTC	357
QY	126	HisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIlePro	145
DB	358	CATTCTAAGTATGATAGTGTGATTTGGTTCATCCCAAGTATGACAGTGTGAAATCCCT	417
QY	146	ArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSer	165
DB	418	CGATATATATCCATCCCTCCTCAGCAGCAGCATCTCAGTGAATTCCTGGAGCGTCC	477
QY	166	ProAspHisHisMetSerProThrGlyAsnIleSerArgAla---ProPhePro	184
DB	478	CCTGAT---CATATGATGCTCCCGTTGGGAACATTTGGCAGACGTGGGCATCCATTTCC	534
QY	185	TyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValala	204
DB	535	TATGTGAACCATTCACCAAAACCCATCAAGGAGTCTTCTGGTAGCCCTGGCAATGTGCA	594
QY	205	TrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThr	224
DB	595	TGGAAAGAGAGATGATGCTGGAAATGAG---GATAAAGGTGCATTTCCCATGGCT	651
QY	225	AsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThr	244
DB	652	AATGCTACTAGCATTTGCTCTTCAGAGGTCGTGGGGTTGGTGACATTTGATGCTACT	711
QY	245	GluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLys	264
DB	712	GACTATAACATGGAGCGCCTTACTGAATGATGAACACGCCAACCTCTATCGAGAAA	771

QY	265	ValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuVal	284
DB	772	GTGCCAAATTTCTGTCATCCAGATAAAATCGTACAGATGGTTATTTGTGTCGCGATTGTT	831
QY	285	ValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeu	304
DB	832	GTCTCTATGATATTTCTGCACTACCGTATACAAATCTCTGTGGTAAATGATACCCGCTA	891
QY	305	TrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPhe	324
DB	892	TGTTGTCTCTCTGTCATATGTCAGATTGGTTTGTCTATCTCTGATCTCGGATCAGTTC	951
QY	325	ProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAsp	344
DB	952	CCCAAGTGGTCCCAATCAACCGTGAATCTTACCTGGATAGCTGGCTTTAAGGTATGAC	1011
QY	345	ArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspPro	364
DB	1012	CGGGAAGGTGAACCATCTCAATTTGGCTCTCTGTGATGTTTGTCTAGTACTGTGGATCCT	1071
QY	365	LeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyr	384
DB	1072	ATGAAGAACCTCTCTGTCCTGCTCACTGCCAATATCTGCTCTCTCCATCTCTGCTGGATTAC	1131
QY	385	ProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAsp	404
DB	1132	CCTGTTGATAAGGTATCTTGTCTATGTCGTATGATGGAGCTGCAATGTTAAATTTGAT	1191
QY	405	AlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAsp	424
DB	1192	GGCTTGTCTGAGACTTCAGAAATTTGCTAGAAATGGTACCGTCTCTGTAAGAGTACAGC	1251
QY	425	IleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLys	444
DB	1252	ATAGAACCCAGAGCTCCAGAGTGTACTTTGGCACAGAAATTTGATTACTTGAAGACAAA	1311
QY	445	ValGlnProSerPheValLysAspArgAlaMetLysArgGlnTyrGluGluPheLys	464
DB	1312	GTTTCAGCTCTTTTGTAAAGATCTGCTGTCATGAAGAGGAATATGAAGAAATTTAA	1371
QY	465	IleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMet	484
DB	1372	GTTCTGTTAATGCTCTGTCGCCAAGGCACAGAAAGTTCTCTGAGGAGGATGGATTATG	1431
QY	485	GlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal	504
DB	1432	CAGATGGCACACTTGGCTGTGTAACAATACACAGAGACCATCTCTGGAATGATTCAGGT	1491
QY	505	PheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyr	524
DB	1492	TTCTTGGTCAATAGTGGAGGCTTGATACCAAGGCAATAGGCTTCCTCTGTTAGTCTAT	1551
QY	525	ValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAla	544
DB	1552	GTGTCTCGTGAGAAACGCTCTGGATTCACACCATTAAGAGGCTGTGTATGAATGCA	1611
QY	545	LeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp	564
DB	1612	CTTGTCTGTATCAGCTGCTCTTACTAATGGAACAATCTGTTGTAATCTTGAATGAT	1671
QY	565	HisTyrIleAsnAsnSerLysAlaValArgGlnAlaMetCysPheLeuMetAspProAsn	584
DB	1672	CATCATCATCAACATAGCAAGCTCTCCGAGAGGCTATGTGCTCTCTTATGATGATCAAC	1731
QY	585	LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn	604
DB	1732	CTAGGAAGGCTGTCTGTATGTCATTTCCCTCAGAGGTTTGGCGGTATCGATAGAAT	1791
QY	605	AspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGly	624
DB	1792	GATCGATATGCAACACAGCAACCGTGTGTTTTCGATATTAATTTGAGAGGCTCTCGATGT	1851
QY	625	IleGlnGlyProValTyrValGlyThrCysValPheAsnArgThrAlaIleTyrGly	644

Db	1852	CTCCAGGACCGATTATGTGGGAACCGGTTGTGTGTTCAACAGAACACGCTCTTTATGGT	1911
QY	645	TyrGluProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLys	664
Db	1912	TATGAACCCCATTAAGCAGAGAGCCAGGTACTTCTTGTGTTTGTGGGGAGCA	1971
QY	665	LysLysAlaSerLysSerLysLysArgSerAspLysLysLysSerAsnLysHisVal	684
Db	1972	AAGAAGACAAAAGATCGAAGCAGAGACGCGAAGAAAAAGTACACAAACATGTG	2031
QY	685	AspSerSerValProValPheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGly	704
Db	2032	GACAGTTCTGTGCCAGTTTTTAATCTTGAAGATATAGAGGAAGGATGAAGGTCTCGA	2091
QY	705	PheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGln	724
Db	2092	TTTGATGATGAGAAATCACTACTCATGTCTCAATAGAGCTTAGAGAAAAAGATTGGTCAA	2151
QY	725	SerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThr	744
Db	2152	TCTATGTGTTTTGTAGCCTCCACTCTGATGGAATATGGTGTGTCTCTCAATCGCAACT	2211
QY	745	ProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSer	764
Db	2212	CCAGATCTCTCTGAAAGAGGCCATACATGTTATCAGCTGTGCTATGAAGACAAAAGT	2271
QY	765	GluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGly	784
Db	2272	GACTGGGGAACCTGAGATTGGATGGATCTATGCTTCTGTTACAGAGATATTCACCTGGA	2331
QY	785	PheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPhe	804
Db	2332	TTCAAGATGATGACAGCTGGCTGGCGTCAATCTACTGCTGCTTAAGCAGCAGCATTC	2391
QY	805	LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeu	824
Db	2392	AAGGGGCTGCCCCAATTAACTTTTCGGATGCTCTGAACACAGGTGCTTGGTGGGCTCTT	2451
QY	825	GlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArg	844
Db	2452	GGTCTCTGGAAATCTTTTTAGCAGCATGTGCCATATGGTATGGTACGGAGACCG	2511
QY	845	LeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeu	864
Db	2512	CTTAAAGTTCCTGGAGAGATTGCTTACATCAACACCACTATTATTCCATCATCGATT	2571
QY	865	ProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMet	884
Db	2572	CCATCTCTCTGTACTGCATATTGCCAGCTATCTGCTCTCTCACTGGGAAGTTCATCA	2631
QY	885	ProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAla	904
Db	2632	CCAGAGATTAGCAACTTTGGCAGTATTGGTTATCTCTCTTTCTGTCATCTTTGCT	2691
QY	905	ThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrpTrpArgAsnGlu	924
Db	2692	ACTGGTATCCTTGAGATGAGTGGAGTGTGTGGCATCGATCAATGGTGGAGGAATGAG	2751
QY	925	GlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeu	944
Db	2752	CAGTTTGGGTTATTGGAGTATTTCGGTCACTATTATTTGGCGTCTTCCAGGGTCTGCTG	2811
QY	945	LysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGlu	964
Db	2812	AAGGTGCTTGCTGGAAATTGATACCACTGCTACCTGTCACCTCAAGAGGCTCTCATGA	2871
QY	965	GlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThr	984
Db	2872	GGCGAATTGCTGAGCTACATGTTCAAGTGGACAACTTCTGATCCACCGACCACT	2931
QY	985	IleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGly	1004

Db	2932	ATTTTGATCATCAACCTGGTCGGTGTGTTGCTGGTATCTCTCAAGCTATCAACAGTGGC	2999
Qy	1005	TyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHis	1024
Db	2992	TACCAAGTCATGGGGTCCGCTCTTTGGGAAGCTCTCTTTGCCTTCTGGGTGATGTCCAC	3051
Qy	1025	LeuTyrProPheLeuLysGlyLeuMetGlyValArgGlnAsnArgThrProThrIleValIle	1044
Db	3052	CTGTACCCCTTCCTTAAGGGTCTCATGGGTGCGGAGAACCGTACACGACCATGTGTGT	3111
Qy	1045	ValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProphe	1064
Db	3112	GTTTGGGCATCTCTTGCATGATCTTCTCTCTGCTGGGTTCGTATCGATCATTC	3171
Qy	1065	ThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys	1080
Db	3172	ACCACCGGTGCACAGCCCTGTATACCCAAAATGTGTATCAACTGC	3219
RESULT 7			
AAZ99533			
ID	AAZ99533	standard; DNA; 3704 BP.	
XX	AC	AAZ99533;	
XX	XX	03-JUL-2000 (first entry)	
DT	DT	DNA encoding a maize cellulose synthase.	
DE	DE	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;	
XX	XX	transgenic plant; plant breeding marker; ss.	
KW	KW	Zea mays.	
OS	OS		
XX	XX	Key	Location/Qualifiers
FT	FT	CDS	272..3496
FT	FT		/*tag= a
FT	FT		/product= "cellulose synthase"
FT	FT		/transl_except= (pos: 1544..1546, aa: Xaa)
FT	FT		/note= "no termination codon given; Xaa is an unspecified
XX	XX		amino acid"
XX	XX	WO200009706-A2.	
XX	XX		
PD	PD	24-FEB-2000.	
XX	XX		
PF	PF	16-AUG-1999; 99WO-US018760.	
XX	XX		
PR	PR	17-AUG-1998; 98US-0096822P.	
XX	XX		
PA	PA	(PION-) PIONEER HI-BRED INT INC.	
XX	XX		
PI	PI	Drugga KS, Helentjaris TG, Bowen BA, Wang X;	
XX	XX		
DR	DR	WPI; 2000-224343/19.	
DR	DR	P-PSDB; AAY84121.	
XX	XX		
PT	PT	New genes which encode maize cellulose synthase polypeptides in plants	
PT	PT	useful for modulating the expression of cellulose synthase in plants	
PT	PT	and to produce transgenic plants expressing the novel protein.	
XX	XX		
PS	PS	Claim 1; Page 191-196; 119pp; English.	
XX	XX		
CC	CC	The present sequence encodes a maize cellulose synthase polypeptide. The	
CC	CC	cellulose synthase can be used for the improvement of stalk quality for	
CC	CC	improved stand or silage. It also provides an increased concentration of	
CC	CC	cellulose in the pericarp, hardening the kernel and improving its	
CC	CC	handling ability. The sequences are used to produce transgenic plants	
CC	CC	seeds expressing the cellulose synthase. The polynucleotide is used for	
CC	CC	modulating, preferably increasing, the level of the synthase in a plant	
CC	CC	cell. The plants are preferably monocots. The polynucleotide is also used	
CC	CC	as a probe or primer in the detection quantitation or isolation of gene	
CC	CC	transcripts. The probes are useful in detecting deficiencies in the level	
CC	CC	of mRNA in screenings for desired transgenic plant. for detecting	

CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 5143.00 Matches: 947
Percent Similarity: 93.99% Conservative: 69
Best Local Similarity: 87.60% Mismatches: 59
Query Match: 89.01% Indels: 6
DB: 3 Gaps: 6

US-09-900-237A-30 (1-1080) x AA299533 (1-3704)

QY	1	MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln	20
DB	272	ATGGACGGC---GGCGACGCCACGAATTCGGGGAAGCATGTGCGCGGAGGTGTGCCAG	328
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
DB	329	ATCTGGCGGACGGCGTGGCGCACCGCGCGGAGCGGCCTCTTCACCGCGTCGACGTC	388
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla	60
DB	389	TGCGGTCTCCCGTGTGCGGCCCATGCTACGAGTACGAGCGCCAGGACGCCACCGCGG	448
QY	61	CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
DB	449	TGCCCCGACGTGCAAGACTAAGTACAGCGCCACAAAGGGAGGCCACCGATACACGGT	508
QY	81	GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
DB	509	GAATATGAGGATGTGGATGCTGACGATGTGAGTGCATCTACCACTACCAAGCATCTG	568
QY	101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly	120
DB	569	CAGATCAGAAAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGCAAACTCACGTGG	628
QY	121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
DB	629	AGT---GATATTGGCTGGCTAAGTATGACAGCGGTCAAATTTGGGCATGGGAAGTATGAC	685
QY	141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160
DB	686	AGTGGTGAGATCCCTCGTGGATATATCCCGTCACTAACTCATGCCAGATCTCAGGAGAG	745
QY	161	IleProGlyAlaSerProAspHisHisMetSerProThrGlyAsnIleSerArgArg	180
DB	746	ATTCTCTGGAGCTTCCCTCTGAT---CATATGATGTCCTCTGTTGGGAACATGGCAGCGT	802
QY	181	Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer	199
DB	803	GGACATCAATTCCTTATGTAATCAATTCCTCAAACCATCGAGGGAGTTCCTCGGTAGC	862
QY	200	IleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLysGlnAspLysGly	219
DB	863	CTTGGCAATGTGATGGAAGAGAGGGTGGATGGATGGAAAAATGAAG---GATAAAGGT	919
QY	220	AlaTyrProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp	239
DB	920	GCAATTCCTATGACCAATGGACAGCAATGCTCCATCAGAGGGCGGTGGAGTTGCTGAT	979
QY	240	IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln	259

DB	980	ATTGATGCTTCTACTGATTATAACATGGAAGATGCTTACTCAATGATGAACCTCGCAA	1039
QY	260	ProLeuSerArgLysValProIleAlaSerSertyIleAsnProTyrArgMetValIle	279
DB	1040	CCTCTATCTAGAAAGTCCCAATTCCTTCAATCCAGATAAATCCGTACAGATGTGTCA	1099
QY	280	ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg	299
DB	1100	GTGCTACGTTTGGCTGTCTATGTCATATCTTCCGCTACCGTATCACACATCCTGTGAAC	1159
QY	300	AsnAlaTyrProLeuTyrPheLeuSerValIleCysGluIleTyrPheAlaLeuSerTyr	319
DB	1160	AATGATCATATCCACTGTGGCTTTTATCCGTCATATGTGAGATCTGGTTCCTTGTCTGG	1219
QY	320	IleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeu	339
DB	1220	ATTTTGGATCAGTTCCTCCAAAGTGTCTCCCAATCAACCGTGAACATACCTTGATAGCTG	1279
QY	340	AlaLeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheVal	359
DB	1280	GCTTTAAGGTATGACCGAAGGTGAACCATCTCAATTAGCTCCTGTTGATATTTTGTG	1339
QY	360	SerThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIle	379
DB	1340	AGTACTGTGGATCCCAATGAAGAGCTCTCTGTGTCACCTGCAATACTGTGCTTCCATC	1399
QY	380	LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer	399
DB	1400	CTTGCTGTCGATTATCCGCTTGACAAGGTATCTTGCTATGTTTCGGATGATGAGGTGCT	1459
QY	400	MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPhe	419
DB	1460	ATGCTGACTTTTGTATGCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGGTTCGGTTC	1519
QY	420	ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp	439
DB	1520	TGTAAGAAGTACCAATAGAGCTTANGGCCCGGATGGTACTTGTGTCAGAAAAATTGAT	1579
QY	440	TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu	459
DB	1580	TACTTGAAGACAAAAGTTCAAACTCATTTGTGAAGAACGCGCGGCGCATGAAGAGAA	1639
QY	460	TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu	479
DB	1640	TATGAAGAAATCAAAAGTTGATCAATGCTTGTAGCCCAAGGACAAAAGTTCCCGAG	1699
QY	480	GluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisPro	499
DB	1700	GAGGATGGATCATGCAAGATGTTACACCTTGGCTGGGAACAATACTAGGACCATCT	1759
QY	500	GlyMetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeu	519
DB	1760	GGAATGATTGAGTTTCTCGGTTCACAGTGGAGGCTTGACGTTGAAGCAATGAACCTT	1819
QY	520	ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla	539
DB	1820	CCTCGTTGTTGTTTATGTCCTCGTGAACACGCTCTGGATTCACCAATCAAGAGAGCT	1879
QY	540	GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu	559
DB	1880	GGTGCCATGAATGCACTGTTTCTGTGATACAGTGTCTTACTAATGGCAATACATGTTG	1939
QY	560	AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe	579
DB	1940	AATCTTATGTGACCACTACATCAATAATAGCAAGGCTCTTCGAGAGCTATGTGTCTC	1999
QY	580	LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp	599
DB	2000	CTTATGACCCCAACCTTAGGAAGGAATGCTGTTATGTCCTCAATTTCTCAGAGGTTGAT	2059
QY	600	GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu	619
DB	2060	GGTATTCATAGGAATGACCCGATATGCAACACAGGAACACTGTGTTTTTCGATATTAAC	2119


```

QY 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 2120 AGAGGCTTTCAGCGCATTTCAGGGCAGCTTTATGTGGAACTGGTTGTGTGTAAACAGA 2179
QY 640 ThrAlaIleTyrGlyTyrGluProPheIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2180 ACGGCGCTTATATGTTGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCCTCG 2239
QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2240 CTTTGTGGGAGGAGAAAGACGTCAAAAATCTAAGAAG--AGCTCGAAAAGAGAGAG 2296
QY 680 SerAsnLysHisValAspSerValProValPheAsnLeuGluAspIleGlnGluGly 699
Db 2297 TCACATAGACAGCGACAGATCTCTACAGATTTTATCTCGAAGATATAGAGAGGG 2356
QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2357 ATTGAAGGTTCTCAGTTTGTATGATGAGAAATCGCTGATTATGCTCAATGAGCTTGGAG 2416
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2417 AAGAGATTGGCCGCTCAGGCTTTTGTAGGCTCTACTCTGATGGAATATGTTGCTGTT 2476
QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2477 CCACAATCTGCAACTCCAGAGCTCTCTCTGAAGAAGCTATTCAATGATCACTAGCTGTGCG 2536
QY 760 TyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db 2537 TATGAGGACAAACTGACTGGGAACTGAGATGGTGGATCTATGTTTCTGTTACAGAA 2596
QY 780 AspileLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2597 GACATTCTCACCGGATTCAGATGATGCTCGAGCTCGCGATCAATCTACTGCTGATGCT 2656
QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2657 AAGCGACACAGCTTTCAGGGATCTGCTCTATCAACCTTCGGATCGTTTGAATCAAGTG 2716
QY 820 LeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyr 839
Db 2717 CTTCGGTGGGCTCTTGGTTCCATTGAAATCTTTTTCAGAGGCAATGTCCTCATATGGTAT 2776
QY 840 GlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2777 GGCTATGAGCCCGGCTTAAATTCCTGGAGAGATTGCTTATATCAACACACAATTTAT 2836
QY 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2837 CCACCTCATCAATCCCGCTCTCTCTGCTACTGATATGCGAGAGTTTGTCTTCTCACT 2896
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleThrPheIleAlaLeuPhe 899
Db 2897 GGGAGGTTTCATCCCAAGAGTAGTAACCTAGAGAGTGTTCGTTTATATGCTCTTT 2956
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrPheSerGlyValGlyIleAspGlu 919
Db 2957 ATCTCAATCTTTCACATGTTATCTTGGATGATAGAGTGGAGTGGTGTGGCATTTGATCAA 3016
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 3017 TGGTGGAGGACAGAGCTTCTGGGTCATTTGGTGGTATTTCTGGCCATTTATTTCGCCGTC 3076
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3077 TTCAGGCTCTCTGAAGGCTTCTGCTGATGACACAGAGCTTCACTGTCACCTCTAAG 3136
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 3137 GCCACTGACGAAGAAGGTGATTTTGCCGAGCTCTACATGTTCAAGTGGACACAGCTTCTG 3196

```

```

QY 980 IleProProThrThrIleLeuIleLeuMetValGlyValValAlaGlyThrSerTyr 999
Db 3197 ATCCCAACCAACACATATTTTGATCATCAACCTGGTCGGCGTGGTCGTCATTTCTCTAC 3256
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3257 GAATCATATAGCGTTTACCGTCATGGGACCTCTTTTCGGGAAGCTCTCTTTGGGTTTC 3316
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3317 TGGGTGATTGTCCACTGTATACCTTCTCAAGSGCTCATGGGAAGCAGAACCCGACG 3376
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3377 CCGACCATGTCTGTTGTCTGGGCTATCTCTTGGTTCGATCTTTTCCCTGATGCGGTT 3436
QY 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3437 CGTATCGATCCATTCACCAACCGGCTCACTGGCCCTGATATCGGAAATGTGGCATCAAC 3496
QY 1080 Cys 1080
Db 3497 TGC 3499

RESULT 8
AAC49550
ID AAC49550 standard; DNA; 3198 BP.
XX
AC AAC49550;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.

```

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144894P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151343P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

```

PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 0 Length: 3198
Score: 4530.00 Matches: 839
Percent Similarity: 87.82% Conservatives: 98
Best Local Similarity: 78.63% Mismatches: 104
Query Match: 78.40% Indels: 26
DB: 3 Gaps: 8

US-09-900-237A-30 (1-1080) x AAC49550 (1-3198)

QY 19 CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys 38
DB 58 TGCCAGATCTGTAGTGACAAATGTGGCAAGACTGTTGATGGAGATCGTTTTGTGGCTGT 117

QY 39 AspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThr 58
DB 118 GATATTGTTCATCCCGCTTGTCCGCTTGTCTAGAGTATGAGAGGAAGATGGGAAT 177

QY 59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArg 78
DB 178 CAATCTTGCTCCTCAGTGCAAAACAGACATACAGAGGCTCAAGAGTAGTCTCTATTCCT 237

QY 79 GlyGluGluGlyAspAspThrAspAlaAspAspGlySer--AspPheAsnTyrProAla 97
DB 238 GGTGATAAAGACGAGGATGGCTTAGCTGATGAAGGTACTGTTGAGTTCACCTACCT-- 294

QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsn 117
DB 295 -----CAGAAGAGAAATTTTCAGAGCGAGTCTTGGTTGGCATCTTACT 339

QY 118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
DB 340 CGTGGGAAGGGAGAGAAATGGGGAAACCCAGTAGTAGATAA----- 381

QY 138 LysTyrAspSerGlyGluIleProArgGlyThrValProSerValThrAsnSerGln--- 156
DB 382 -----GAGGTCTCTCACAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 426

QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
DB 427 ACTTCAGAGAGTTTCTGCTGCCTCACCTGCAAGCGCTCTCTGATCTCTCTACTATGCT 486

QY 177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
DB 487 GGGGGAAGCGCCTTCCTTATCATCAGATGTCATCAATCAATCACCANAATAGAGGATGTG 546

QY 196 PheSerGlySerIleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLys 215
DB 547 GATCCTGTTGGACTCGGGAATGAGCTTGAAGGAGAGAGTTGATGCTCGAANAATGAAG 606

QY 216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
DB 607 CAAGAGAAGAAATACTGTGCTCTGTC-----AGCAGCAGGCTGTCTTCTGAA---AGA 654

QY 236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
DB 655 GGTGGAGTAGATATTGATGCAGACAGATATCTCAGATAGAGGCTCTGCTGAATGAC 714

QY 256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
DB 715 GAACGAGGAGGAGGCTCTGTCAAGAAAGTTTCAATTCCTTCATCACCGATCATCTCTTAC 774

QY 276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
DB 775 AGAATGTTATTATGTGGGCTTGTATCTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 834

```

```

QY 296 AsnProValArgAsnAlaTyrProLeuTyrPheLeuSerValIleCysGluIleTyrPhe 315
DB 835 AACCCAGTGCCTTGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894

QY 316 AlaLeuSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyr 335
DB 895 GCCTTATCTCTGGATTGTGGATCATGTTTCCCAAGTGGTTTCTGTGAACCGTGAACCTAC 954

QY 336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaIleVal 355
DB 955 CTCGACAGCTTGTCTTAAGATATGATGCTGAAGGTGAGCCATCAAGTTAGTCTGTT 1014

QY 356 AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375
DB 1015 GACATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074

QY 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395
DB 1075 GTGCTCTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134

QY 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
DB 1135 GATGCTGCTGCTATGCTTATCATTTGCAATCATTGCAAAACATCAGAGTTTCTCGTAA 1194

QY 416 TrpValProPheValLysTyrAspIleGluProArgAlaProGluLysPheTyrPheCys 435
DB 1195 TGGGTACCATTTTGCAGAAATATAGTAGAGCTCGTGACCAAGATGGTACTTTGCT 1254

QY 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
DB 1255 GCGAAATAGATATTCTTCAAGGATAAAGTTTCAAGATCATTTGTCAGAGATCTAGAGCT 1314

QY 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
DB 1315 ATGAAGAGGGAATAGAGAAATTTAAATCCGAATCAATGACATTTGTTCCAAAGCCCTA 1374

QY 476 LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr 495
DB 1375 AAATGTCTGAAGAGGGTGGTTTATGCAAGATGGCACACCGTGGCTGGAAATATATACA 1434

QY 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
DB 1435 AGGACCATCCAGAAATGATCCAGGTCTTCTTAGGCAAAATGCTGGAGTTGATGCAGAG 1494

QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
DB 1495 GGCAATGAGCTCCCGCTTGTGTATATGTTTCTCGAGAAAGCCAGGATTCAGGAC 1554

QY 536 HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
DB 1555 CACAAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1614

QY 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
DB 1615 CTTTTCATCTGAAATCTTGATTTGATTCATTAATAATAACAGCAAGCTTAAGAGAA 1674

QY 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
DB 1675 GCAATGTCTCTGATGGACCCCAACCTCGGAAGCAAGTTTGTATGTTCAGTTCCCA 1734

QY 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
DB 1735 CAAAGATTTGATGCTATCGATAAGAACGATAGATATGCTAATCGTAATCCGTTGTTCT 1794

QY 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
DB 1795 GATATTAACTGAGAGGTTTAGATGGGATTCAGGACCTGTATATGTCGGAACCTGATGT 1854

QY 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653
DB 1855 GTTTTCAACAGAACATTTATACGGTTATGAACCTCCCAATAAAGTAAACACAGAAG 1914

```



```
CC standardise OS field)
XX
SQ Sequence 3614 BP; 963 A; 717 C; 863 G; 1071 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3614
Score: 4499.00 Matches: 835
Percent Similarity: 87.44% Conservative: 98
Best Local Similarity: 78.26% Mismatches: 108
Query Match: 77.86% Indels: 26
DB: 2 Gaps: 8

US-09-900-237A-30 (1-1080) x AAV06567 (1-3614)

QY 19 CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys 38
Db 274 TGCAGATCTGTAGTCAATGTCAGACACTGTGTAGTGGAGATCGTTTGTGGCTTGT 333
QY 39 AspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThr 58
Db 334 GATATTGTGTCAATTCCTCAGTTTGTGGCTTGTGTACGAGTATGAGAGGAAGATGGGAAT 393
QY 59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArg 78
Db 394 CAATCTGTCTCAGTGCAGAAACCAAGATACAGAGGCTCAAGAGGTAGTCTGTATTCCT 453
QY 79 GlyGluGlyAspAspThrAspAlaAspAspGlySer--AspPheAsnTyrProAla 97
Db 454 GGTGATAAAGACGAGGATGGCTTAGCTGATGAGGTACTGTGAGTTCAACTACCCCT--- 510
QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
Db 511 -----CAGAAGGAGAAATTTTCAGAGCGGATCTTGTTGGTGGCATCTTACT 555
QY 118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
Db 556 CGTGGGAAGGAGAGAAATGGGGAAACCCAGTATGATAA----- 597
QY 138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
Db 598 -----GAGGTCTCTCACAATCATCTTCTCTCTCTCTCAGCAGCAGACAGAT 642
QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
Db 643 ACTTCAGAGAGATTTCCTGCTGCTCCTCAGTGAAGCGCTCTCTGTATCTTACTATCGCT 702
QY 177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
Db 703 GGGGGAAGCGCTTCTCTATTCATCAGATGTCATCAATCAATCACCAMAATAGAAGATTGTG 762
QY 196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLys 215
Db 763 GATCCTGTTGGACTCGGAATGTAGCTTGAAGAGAGAGTGTATGCTGGAAATGAAG 822
QY 216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
Db 823 CAAGAGAAGAATACTGCTCTCTGTC-----AGCAGCGAGGCTGCTTCTGAA--AGA 870
QY 236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
Db 871 GTTGGAGTAGATATTGATGCCAGCACAGATATCTTAGCAGATGAGGCTCTGCTGAATGAC 930
QY 256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
Db 931 GAAGCGAGGCGAGCTTCTGTCAAGAAAGTTTCAATCTCTTCATCAGCATCAATPCTTAC 990
QY 276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
Db 991 AGAATGGTATTATGTGCGGCTTGTATCTCTTGTCTCTTCTGCTATACCGTATAACA 1050
QY 296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe 315
Db 1051 AACCCAGTGCCAATGCCCTTGTCTGTATGCTGCTGTATGATATGTGAGATCTGCTT 1110
```

```
QY 316 AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr 335
Db 1111 GCCTTATCTCTGGATTTTGGATCAGTTTCCCAAGTGGTTTCTGTGAACCGGTGAACCTAC 1170
QY 336 LeuAspArgGluAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal 355
Db 1171 CTCGACAGGCTTGTCTTAAAGATATGATCGTGAAGGTGAGCCATCACAGTTAGTCTGTT 1230
QY 356 AspIlePheValSerThrValAspProLeuLysGluProIleValThrAlaAsnThr 375
Db 1231 GACATTTTCTGTGAGTACTGTGTACCTTGAAGAGCCCTTGAAGAGCCCTTGTGACAGCAACACA 1290
QY 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395
Db 1291 GTGCTCTCTATCTGTGCTGTGTACTACCCAGTTGACAGGTGCTCTGTATGTATGTCTGAT 1350
QY 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
Db 1351 GATGGTGTCTGTATGTATCATTTGAATCACTTGCAGAAACATCACAGTTTGTCTCGTAA 1410
QY 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435
Db 1411 TGGTACCATTTTGCAGAAATATAGCATAGAGCCTCGTGCACCAAGATGGTACTTTGCT 1470
QY 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
Db 1471 GCGAAATAGATTACTTGAAGGATTAAGTTCAGACATCATTTGTCAAGATCGTAGAGCT 1530
QY 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
Db 1531 ATGAAGAGGGAATATGAGGAATTTAAATCCGAATCAATGCACTTGTTCCTCAAGGCCCTA 1590
QY 476 LysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr 495
Db 1591 AAATGCTCTGAGAGAGGTGGTGTATGCAAGATGCAACCGTGGCTGGAAATAATACA 1650
QY 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
Db 1651 GGGGACCATCCAGGAATGATCCAGTCTTCTTAGGGCAATTTGGTGGCTGTATGATCCAGAG 1710
QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
Db 1711 GCGAATGAGCTCCCGCTTGTGTATGTTCTCGAGAAAAAGCAGCAGGATTCAGCAC 1770
QY 536 HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
Db 1771 CACAAAAGGCTGCTGTATGATGCACTGTGTGAGAGTTTCAGCAGTTCTTTACCAATGGA 1830
QY 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
Db 1831 CCTTTCATCTTGAATCTTGTATGATCAATTACATAAATAACCAAGCCCTTAAGAGAA 1890
QY 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
Db 1891 GCAATGCTTCTGTATGAGCCCAACCTCGGAAAGCAAGTTTGTATGTTCAGTTCCCA 1950
QY 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
Db 1951 CAAAGATTGATGATCGATCGATAGAACGATAGATATGCTAATCGTAATCCGTTCTTT 2010
QY 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
Db 2011 GATATTAATCTGAGAGGTTTAGATGGATTCAGGACCTGTATATGTCCGGAACCTGATGT 2070
QY 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653
Db 2071 GTTTTCAACAGAACACCATATACGGTTATGAACCTTCAATTAAGTAAACACACAGAG 2130
QY 654 ProGlyPheLeuAsnLeuSerLysGlyLysLysLysAlaSerLysSerLysLysArg 673
Db 2131 CCAAGTCTTTATCTAAGCTCTGTGTGTGATCAAGAAAGAAATTTCCAAAGCTAAGAAA 2190
```

[illegible]

CC with a protein
 XX Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3725
 Score: 3993.50 Matches: 749
 Percent Similarity: 80.64% Conservative: 134
 Best Local Similarity: 68.40% Mismatches: 153
 Query Match: 69.12% Indels: 59
 DB: 3 Gaps: 20

US-09-900-237A-30 (1-1080) x AAZ99500 (1-3725)

QY 2 AspGlyAspAlaAspAla---LeuLysSerGlyArgHisGlyAlaGlyValCysGln 20
 DB 245 GACGGCGAGCCCTGTCGGGCTAAGCCACGAGAGTGCGAATGGGCGAGTGTGCCAG 304
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 305 ATTTGTGGCGACACTGTTGGCGCTTCAGCCACTGGTGATGCTTTGTTGCCCTGCAATGAG 364
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 DB 365 TGTGCTTCCCTGCTGCGCGCCCTTGCTATGATGATGAGTACGAGCGAAGGAGGAACCAATGC 424
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 425 TGCCCTCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTTCATGGTGTAT 484
 QY 81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly 99
 DB 485 GAT---GAGGAGGAAGATGTTGATGACCTGGGACAATGAATCAACTAT-----529
 QY 100 ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly 119
 DB 530 -----AAGCAA-----535
 QY 120 GlySerGlyAsnValGlyHisProLysTyr-----AspSerGlyGluIleGly 135
 DB 536 -----GSCATGGGAAGGGCCGAGGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCT 589
 QY 136 LeuSer---LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsn 154
 DB 590 TCATCTGTCTGCCATGAC-----CCACACCATCGAATTCACGCGCTTACAACT 637
 QY 155 --SerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis---MetMetSer 172
 DB 638 GGACACAGATATCTGGAGAGATCCCTGATCCCTGACCTCCCGTCACTTCTATCCGCACT 697
 QY 173 ProThrGlyAsn---IleSerArgArgAlaProPheProTyrValAsnHisSerProAsn 191
 DB 698 CCAACATCGAGCTATGTTGATCCAAAGCGTTCAGTTCCTGTG-----AGGATTTGGGAC 751
 QY 192 ProSerArgGluPhe---SerGlySerIleGlyAsnValAlaTrpLysGluArgValAsp 210
 DB 752 CCTCGAAGGACTTGAATTCCTATGGGCTTAATAGTGTGATCGAAGGAAGATTCAG 811
 QY 211 GlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAla 230
 DB 812 AGCTGGAGGGTTAAACAGGACAAAATAATGTTGCAAGTGAATA-----856
 QY 231 ProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAsp 250
 DB 857 ---AAATATCCAGGGCTAGAGGAGCATGGAG---GGGACTGGCTCAAAATGGAGAAGAT 910
 QY 251 AlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIleAlaSerSer 270
 DB 911 ATGCAAAATGGTTGATGATGACGCGCTACCTTGGCCGCGCATTTGGCCAAATTTCTCTAAC 970
 QY 271 LysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIlePheLeu 290
 DB 971 CAGCTCAACCTTTACCGGATAGTAATCATCTCCGCTTATCATCTCTGCTGCTCTCTCTTC 1030

QY 291 HisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIle 310
 DB 1031 CAATATCGTATCAGTCTCAGTCGCTAATGCTTATGGATTGGCTAGTATCTGTTATC 1090
 QY 311 CysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIle 330
 DB 1091 TGTGAGTCTGGTGTGGCTTGTCTGGCTTCTAGATCAGTCCCAAAATGGTATCCAATC 1150
 QY 331 AsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSer 350
 DB 1151 AACCGTGAGACATACTCGACAGGCTTGATGAGGTATGATAGAGGAGGAGGCCATCA 1210
 QY 351 GlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProIle 370
 DB 1211 CAGCTGGCTCCCATTTGATGATGCTTTGTGTCAGTACAGTGGATCCATTGAAGGAACCTCCACTG 1270
 QY 371 ValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSer 390
 DB 1271 ATCAGAGCCACACACTGTTTGTCCATTCTGTGTGGATTACCTGTTGACAAAGTGTC 1330
 QY 391 CysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSer 410
 DB 1331 TGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGAGTCTCTCTGAAACTGCC 1390
 QY 411 GluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArgAlaPro 430
 DB 1391 GAATTTGCTAGAAAGTGGGTTCCTTTTGAAGAAGACAAATATTGAACCAAGAGCTCCA 1450
 QY 431 GluPheTyrPheCysGlnLysIleAspTyrLysAspLysValGlnProSerPheVal 450
 DB 1451 GAATTTTACTTTGCTCAAAAATAGATTACCTCGAAGGACAAAATCAATGCCCTT 1510
 QY 451 LysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeu 470
 DB 1511 AAGGMAAGACCGAGCAATGAAGAGAGAGTATGAGAAATTCAAAATAAGAAATCAATGCCCTT 1570
 QY 471 ValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrp 490
 DB 1571 GTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGTGGACCATGGCTGATGGAACCTGCTGG 1630
 QY 491 ProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGly 510
 DB 1631 CCTGGGAATAACCTTAGGACCATCTCGGCATGATTCAGGTGTCTTCTGGGCGACAGTGGT 1690
 QY 511 GlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArg 530
 DB 1691 GGGCTTGACACTGATGGAATGAATTAACCACTCTTGTCTATGCTCTTCGTGAAAGAGA 1750
 QY 531 ProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAla 550
 DB 1751 CCAGGCTTTGAGCATCAAGAGGCTGGTGCATGAATGAATGACCTGATTCGTGTATCTGCT 1810
 QY 551 ValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSer 570
 DB 1811 GTGCTGACAAATGCTGCTATCTTCTCAATGTGGATTGTGACCATTACTTCAATAGCAGC 1870
 QY 571 LysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCys 590
 DB 1871 AAAGCTCTTAGAAGAGCAATGTGCTTCATGATGGATCCAGCTCTAGGAAGAAACTTGT 1930
 QY 591 TyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArg 610
 DB 1931 TATGTACAAATTTCCACAAAGATTGTGTCATTGACCTTGACGATCGATGCTAATAGG 1990
 QY 611 AsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
 DB 1991 AACATGCTCTCTTGTGATATCAACATGAAGGCTGTAGATGGCATTCAGGGTCCAGTCTAT 2050
 QY 631 ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLys 650
 DB 2051 GTGGGAACAGGATGCTGTTCAATAGCAGGCTTTGTATGGATGATGATGCTCTGTTTGA 2110

QY	651	AlaLys-----LysProGlyPheLeuAlaSerLeuCyCysGlyGlyLeuLysLysAlaSer	668
DB	2111	GAAGCTGATCTGGAACCTAACATTGTTTAAGAGCTGCTGTGTAGAGGAAG---AGA	2167
QY	669	LysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValasSerSerVal	688
DB	2168	AAGAACAAGAGTTATATGATGATGCTCAAGCCGCTATTATGAGAGAGAACAGAAATCTTCAGCT	2227
QY	689	ProValPheAsnLeuGluAspIleGluGlyValGluGlyAlaGlyPheAspAspGlu	708
DB	2228	CCCATCTTTTAACATGGAAGACATCGAGGAGGGTATTGAA-----GGTTATGAGAGTAA	2281
QY	709	LysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPhe	728
DB	2282	AGTTCAGTGTATGTCGCCAGAGAAATTGGAGAAACGCTTTGGTCAGTCTCCAATCTTC	2341
QY	729	ValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeu	748
DB	2342	ATTGCATCCACCTTTATGACTCAAGGTGGGCATACCACTTCAACAAACCAGCTTCTCTA	2401
QY	749	LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysAspSerGluTyrGlyThr	768
DB	2402	CTGAAGGAAGCATATCCATGTTATCAGCTGTGGGTACGAGACAAACTGAATGGGAAAA	2461
QY	769	GluIleGlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHis	788
DB	2462	GAGATTGGCTGGATCTATGTTTCAGTTACAGAGCATATTCTGACTGGGTTTAAATATGCAT	2521
QY	789	AlaArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla	808
DB	2522	GCAAGAGGCTGGCAATCAATCTACTGATGATGCCACCAAGACCTTGTTCGAAGGGTCTGCA	2581
QY	809	ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGlu	828
DB	2582	CCAAATCAATCTTCTGATGCTCTTAATCAGTGCTCCGTTGGGCTCTTGGGTCACTGGAA	2641
QY	829	IleLeuPheSerArgHisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeu	848
DB	2642	ATTCTGCTTAGCAGACATTTGCTCTATATGTTATGGCTTACAATGGCGAATGAAGCTTTTG	2701
QY	849	GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuVal	868
DB	2702	GAGAGGCTGGGCTTACATTAACACCATTTGTTATCCAAATCACTCTGTTCGCTTATTCGCC	2761
QY	869	TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSer	888
DB	2762	TATTTGTGCTCTCCGCTATCTGCTCTTACCAATAAATTATCATCTCCTGAGATTAGT	2821
QY	889	AsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu	908
DB	2822	AATTATGCTGGAAATGTTCTTCATCTCTTTTGGCTCCCATTTTTCGCACTGTTATATTG	2881
QY	909	GluMetArgTyrSerGlyValGlyIleAspGluTyrTyrArgAsnGluGlnPheTyrVal	928
DB	2882	GAGCTCAGATGAGTGGTGTGGCAATGAGATTGGTGGAGAAATGACGAGTTTGGGTT	2941
QY	929	IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla	948
DB	2942	ATTGTGGGCACCTCTGCCCATCTCTTCGGGGTGTCCAGGGTCTGCTGAAAGTGTGGCT	3001
QY	949	GlyIleAspThrAsnPheThrValThrSerLysAlaIleAsnAspGluGlyAspPheAla	968
DB	3002	GGGATTGATACCAACTTACAGTTACCTCAAAGGCATCTGATGAGGATGGCAGCTTGTCT	3061
QY	969	GluLeuTyrMetPheLysThrThrThrLeuLeuIleProProThrThrIleLeuIleIle	988
DB	3062	GAGCTATATGTTTCAAGTGGACCAAGTTTGTCTATCCCTCCGACCACTGTTCTGTCTATT	3121
QY	989	AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyr	1008
DB	3122	AACCTGGTGGAAATGTTGGCAGGAATTTGTTATGTCATTAAACAGCGGTACCAATCTCTGG	3181
QY	1009	GlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPhe	1028

Db	3182	GGTCCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTGATCTCCATCTCTACCCCTTC	3241
Qy	1029	LeuIysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTrpAlaVal	1048
Db	3242	CTCAAGGCTCTCATGGGCAGGCAGACCGCACGCAACATCGTCATCGTTGGTCCATC	3301
Qy	1049	LeuLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProPhe-----Thr	1065
Db	3302	CTCCTTGGCTATCTTCTCTCTGCTGGGTGAAGATCGATCTTTCATCTCCCGACA	3361
Qy	1066	ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys	1080
Db	3362	CAGAAAGCTCCGCC-----TTGGGCAATGTGTGTAAGTGC	3400
RESULT 11			
AAZ99521			
ID	AAZ99521	standard; DNA; 3725 BP.	
XX			
AC	AAZ99521;		
XX			
DT	03-JUL-2000	(first entry)	
XX			
DE		DNA encoding a maize cellulose synthase.	
XX			
KW		Maize; cellulose synthase; stalk quality; stand; silage; cellulose;	
KW		transgenic plant; plant breeding marker; ss.	
XX			
OS		Zea mays.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	179..3397	
FT		/*tag= a	
FT		/product= "cellulose synthase"	
FT		/note= "no termination codon given"	
XX			
PN	WO200009706-A2.		
XX			
PD	24-FEB-2000.		
XX			
PF	16-AUG-1999;	99WO-US018760.	
XX			
PR	17-AUG-1998;	98US-0096822P.	
XX			
PA	(PION-) PIONEER HI-BRED INT INC.		
PI	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;		
XX			
DR	WPI; 2000-224343/19.		
DR	P-PSDB; AAY84117.		
XX			
PT	New genes which encode maize cellulose synthase polypeptides in plants		
PT	useful for modulating the expression of cellulose synthase in plants and		
PT	to produce transgenic plants expressing the novel protein.		
XX			
PS	Claim 1; Page 160-165; 119pp; English.		
XX			
CC	The present sequence encodes a maize cellulose synthase polypeptide. The		
CC	cellulose synthase can be used for the improvement of stalk quality for		
CC	improved stand or silage. It also provides an increased concentration of		
CC	cellulose in the pericarp, hardening the kernel and improving its		
CC	handling ability. The sequences are used to produce transgenic plants and		
CC	seeds expressing the cellulose synthase. The polynucleotide is used for		
CC	modulating, preferably increasing, the level of the synthase in a plant		
CC	cell. The plants are preferably monocots. The polynucleotide is also used		
CC	as a probe or primer in the detection quantitation or isolation of gene		
CC	transcripts. The probes are useful in detecting deficiencies in the level		
CC	of mRNA in screenings for desired transgenic plant, for detecting		
CC	mutations in the gene, for monitoring upregulation of expression or		
CC	changes in enzyme activity in screening assays of compounds, for		
CC	detection of any number of allelic variants of the gene, or for use as		
CC	molecular markers in plant breeding programs. The isolated nucleic acids		
CC	of the present invention can also be used for recombinant expression of		

CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein

XX
 SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3725
 Score: 3993.50 Matches: 749
 Percent Similarity: 80.64% Conservative: 134
 Best Local Similarity: 68.40% Mismatches: 153
 Query Match: 69.12% Indels: 59
 DB: 3 Gaps: 20

US-09-900-237a-30 (1-1080) x AA299521 (1-3725)

QY 2 AspGlyAspAlaAspAla---LeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 |||||
 DB 245 GACGGCGACGCGCGCTGTCCCGGCTAAGCCACGAGAGTGCGAATGGCGAGGTCTGCCAG 304
 |||||
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 |||||
 DB 305 ATTGTGGCGACACTGTGTGGCGTTTCAGCCACTGGTGATGTCTTTGTTGCTGCAATGAG 364
 |||||
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 |||||
 DB 365 TGTGCTTCCCTGTCTGCGCGCTTGCTATGATGATGAGTAGCGGCGAGGAAGGAACCAATGC 424
 |||||
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 |||||
 DB 425 TGCCCTCAGTCAAGACTAGATACAGACACAGAAAGTAGGCTCGAGTTCATGCTGAT 484
 |||||
 QY 81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly 99
 |||||
 DB 485 GAT---GAGGAGGAGAGATGTTGATGACCTGGGCAATGAATTCACATAT--- 529
 |||||
 QY 100 ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly 119
 |||||
 DB 530 -----AAGCAA----- 535
 |||||
 QY 120 GlySerGlyAsnValGlyHisProLysTyr-----AspSerGlyGluIleGly 135
 |||||
 DB 536 -----GGCAATGGAGAGGCCAGAGTGGCAGCTTCAAGAGAGATGACGCTGATCTGCT 589
 |||||
 QY 136 LeuSer---LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsn 154
 |||||
 DB 590 TCATCTGCTGCCATGAC-----CCACACCATCGGATTCACACGCTTACAAAGT 637
 |||||
 QY 155 ---SerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis---MetMetSer 172
 |||||
 DB 638 GGACACAGATATCTGAGAGATCCCTGATGCATCCCTGACCGTCACTTCTATCCGCAGT 697
 |||||
 QY 173 ProThrGlyAsn---IleSerArgArgAlaProPheProTyrValAsnHisSerProAsn 191
 |||||
 DB 698 CCAACATCGAGCTATGTTGATCCCAAGCGTTCCAGTCTCTCTG-----AGGATTCGAC 751
 |||||
 QY 192 ProSerArgGluPhe---SerGlySerIleCysAsnValAlaTrpLysGluArgValAsp 210
 |||||
 DB 752 CCTCGAAGAGCTTGAATCTCTATGGCGCTTAATAGTGTGACTGGAAGGAAGAGTTGAG 811
 |||||
 QY 211 GlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAla 230
 |||||
 DB 812 AGCTGGAGGGTTAAACAGGACAAATAATGTTGCAAGTGCAATAT----- 856
 |||||
 QY 231 ProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAsp 250
 |||||
 DB 857 ---AAATATCCAGAGGCTAGAGGAGACATGGAG---GGGACTGGCTCAATGGAGAGAT 910
 |||||
 QY 251 AlaLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIleAlaSerSer 270
 |||||
 DB 911 ATGCAAATGTTGATGATGACCGCTACCTTTGAGCGGCATTTGTGCCAATTTCTCTCAAC 970
 |||||

QY 271 LysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIlePheLeu 290
 |||||
 DB 971 CAGCTCAACGCTTTACCGGATAGTAATCAATCTCCGCTTATCATCTGTGCTTCTTCTTC 1030
 |||||
 QY 291 HisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIle 310
 |||||
 DB 1031 CAATATCGTATCAGTCATCCAGTGGCGTAATGCTTATGGATTGTGGTAGTATCTGTATC 1090
 |||||
 QY 311 CysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIle 330
 |||||
 DB 1091 TGTGAGTCTGCTGCTTGTCTGGCTTCTAGATCAGTTCCTCAAAATGGTATCCAAATC 1150
 |||||
 QY 331 AsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSer 350
 |||||
 DB 1151 AACCGTGGACATATCTCGACAGCGCTTCATTCAGGTATGATAGAGAGGAGGAGCCATCA 1210
 |||||
 QY 351 GlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProIle 370
 |||||
 DB 1211 CAGCTGGCTCCCATTTGATGATGCTTTGTTCAGTACAGTGGATCCATTCGAAGAACCTCCACTG 1270
 |||||
 QY 371 ValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSer 390
 |||||
 DB 1271 ATCACAGCCAAACACTGTTTGTCCATCTTGTCTGTGGATTACCTGTTGACAAAGTGTCA 1330
 |||||
 QY 391 CysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSer 410
 |||||
 DB 1331 TGCTATGTTCTGATGATGGCTCAGCTATGCTGACTTTTGTAGTCTCTCTCTGAAACTGCC 1390
 |||||
 QY 411 GluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArgAlaPro 430
 |||||
 DB 1391 GAATTTGCTAGAAAGTGGGTTCCTTTTGTAAAGAACACAAATATTGAACCAAGAGCTCCA 1450
 |||||
 QY 431 GluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSerPheVal 450
 |||||
 DB 1451 GAATTTTACTTTGCTCAAAAATAGATTACTCTGAAGGACAAAATTCACACCTTCATTTGT 1510
 |||||
 QY 451 LysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeu 470
 |||||
 DB 1511 AAGGAAGACGAGCAATGAAGAGAGAGATGAAGAATTCAAAATAAGAATCAATGCCCTT 1570
 |||||
 QY 471 ValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrp 490
 |||||
 DB 1571 GTTCCAAAGCACAGAAAGTCCCTGAAGAGGGTGGACCATGGCTGATGGAACTCTTGG 1630
 |||||
 QY 491 ProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGly 510
 |||||
 DB 1631 CCTGGGAATAACCTAGGACCATCTCTGGCATGATTCAGGTGTTCTTGGGACACAGTGGT 1690
 |||||
 QY 511 GlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArg 530
 |||||
 DB 1691 GGGCTTGACACTGATGGAAATGAATTACCACGTCTTGTCTATGCTCTCTCGTGAAGAGAGA 1750
 |||||
 QY 531 ProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAla 550
 |||||
 DB 1751 CCAGGCTTTTCAGCATCAAGAGCGCTGGTGCATGAATGATGCATGATTCTGTGTATGCT 1810
 |||||
 QY 551 ValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSer 570
 |||||
 DB 1811 GTGCTGACAAATGTTGCTCTCTCAATGTGGATTGTGACCATTTACTTCAATAGCAGC 1870
 |||||
 QY 571 LysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCys 590
 |||||
 DB 1871 AAAAGCTTTAGAGAAAGCAATGTGCTTCATCGATCCAGCTCTAGGAGAGGAAACTTGT 1930
 |||||
 QY 591 TyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArg 610
 |||||
 DB 1931 TATGTCAATTTCCAAAGATTGTGTCATTTGACCTTGACATTCGATATGCTAATAGG 1990
 |||||
 QY 611 AsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
 |||||
 DB 1991 AACATAGTCTCTTTTGTATATCAACATGAAAGGCTTAGATGGCAATTCAGGTCCAGTCTAT 2050
 |||||

```
Qy 631 ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProPheLeuLys 650
Db 2051 GTGGAAACAGAGATGCTGTTTCAATAGGAGGCTTTGTATGGATATGATCCTCTTTTGACT 2110
Qy 651 Alalys-----LysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSer 668
Db 2111 GAAGCTGATCTGAACCTTAACATTTGTTAGAGCTGCTGGTAGAAGGAAG--AGA 2167
Qy 669 LysSerLysLysArgSerSerAspLysLysSerAsnLysHisValAspSerSerVal 688
Db 2168 AAGAACAAAGAGTTATATGATAGTCAAGCCGTAATTATGAGAGAACAGAAATCTTCAGCT 2227
Qy 689 ProValPheAsnLeuGluAspIleGluGlyValGluGlyAlaGlyPheAspAspGlu 708
Db 2228 CCATCTCTTAAATGAACATCGAGAGGGGTAITGAA-----GGTTATGAGGATGAA 2281
Qy 709 LysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPhe 728
Db 2282 AGGTGAGTCTTATGTCACAGAGAAATTTGGAGAAACGCTTTGGTCAGTCTCAATCTTC 2341
Qy 729 ValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeu 748
Db 2342 ATTGCATCCACCTTTATGACTCAAGGTGGCATAACCACTTCAACAAACCCAGCTTCTCTA 2401
Qy 749 LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThr 768
Db 2402 CTGAAGGAAGCTTATCATGTTATCAGCTGTGGGTACGAGGACAAACTGATGGGGGAAA 2461
Qy 769 GluIleGlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHis 788
Db 2462 GAGATTGGCTGATCTATGTTTCAGTTCAGAGGATATTCTGACTGGGTTTAAATGTCAT 2521
Qy 789 AlaArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla 808
Db 2522 GCAAGAGGCTGGCAATCAATCTACTGCATGCCACCAACGCTTGTTCAGAGGGTCTTGCA 2581
Qy 809 ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGlu 828
Db 2582 CCAATCAATCTTCTGATCGCTTAATCAGGTGCTCGTTGGGCTCTTGGGTGAGTGAA 2641
Qy 829 IleLeuPheSerArgHisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeu 848
Db 2642 ATTCTGCTAGACAGACATTCCTATATGATGCTACAAATGCGGCGATTTGAAGCTTTTG 2701
Qy 849 GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuVal 868
Db 2702 GAGAGGCTGGCTTACATTAACACCATTTGTTTATCCATCATCATCTGTCGCTTATCGCC 2761
Qy 869 TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSer 888
Db 2762 TATTGTGTGCTTCTGCTATCTGCTTCTTACCAATTAATTTATCATTCCTGAGATTAGT 2821
Qy 889 AsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu 908
Db 2822 AATTATGCTGAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2881
Qy 909 GluMetArgTyrSerGlyValGlyLysAspLeuTyrPheArgAsnGluGlnPheTyrVal 928
Db 2882 GAGCTCAGATGGAGTGTGTTGGCATTTGAAGATTGGTGGAGAAATGAGCAGTTTGGGT 2941
Qy 929 IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAla 948
Db 2942 ATTGGTGCACCTTGGCCATCTCTTGGCGGTGTCCAGGCTCTGCTGAAAGTGTGGCT 3001
Qy 949 GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAla 968
Db 3002 GGGATTGATACCACTTACAGTTACCTCAAGGATCTGATGAGGATGGGCACTTTGCT 3061
Qy 969 GluLeuTyrMetPheLysThrThrLeuLeuIleProThrThrThrIleLeuIleIle 988
Db 3062 GAGCTATATGTTTCAAGTGACCATGTTGCTCATCTCCCTCCGACCACTGTTCTTGTCAT 3121
Qy 989 AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrp 1008
```

```
Db 3122 AACCTGGTCGGAATCGTGGCAGGAATTTGATGTCATTAACAGCGGCTACCAATCTCGG 3181
Qy 1009 GlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPhe 1028
Db 3182 GGTCCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTGATCTCCATCTCTACCCCTTC 3241
Qy 1029 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTrpAlaVal 1048
Db 3242 CTCAGGGTCTCATGGCAGGCGAGAACCGACGCCAACCAATGTCATCGTTTGGTCCATC 3301
Qy 1049 LeuLeuAlaSerIlePheSerLeuLeuTyrValArgValAspProphe-----Thr 1065
Db 3302 CTCCTTGGCTCATCTCTCTCTTCTGCTGGGTGAAGATCGATCTTTCATCTCCCGACA 3361
Qy 1066 ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3362 CAGAAAGCTGCCGCC-----TTGGGGCAATGCTGGTGTGAAGTGC 3400
```

RESULT 12

AAZ99506

ID AAZ99506 standard; DNA; 3725 BP.

XX AAZ99506;

AC AAZ99506;

XX 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers

XX CDS 179..3397

XX /tag= a

XX /product= "cellulose synthase"

XX /note= "no termination codon given"

XX WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US018760.

XX 17-AUG-1998; 98US-0096822P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

XX P-PSDB; AAY84112.

New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein.

Claim 1; Page 121-126; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting or mutations in the gene, for monitoring upregulation of expression or

CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein

XX
 SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3725
Score:	3993.50	Matches:	749
Percent Similarity:	80.64%	Conservative:	134
Best Local Similarity:	68.40%	Mismatches:	153
Query Match:	69.12%	Indels:	59
DB:	3	Gaps:	20

US-09-900-237A-30 (1-1080) x AAZ99506 (1-3725)

QY	2	AspGlyAspAlaaspala---LeuLysSerGlyArgHisGlyAlaGlyAspValCysGln	20
DB	245	GACGGCGAGCGCTGTCCTCCGGCTAAGACCCACGAGAGTCCGAATGGCGAGGTCTGCCAG	304
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysaspval	40
DB	305	ATTGTGGCGACACTGTTGGCGTTTCAGCCACTGGTGATGCTTTGTTCCTCGCAATGAG	364
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla	60
DB	365	TGTGCTTCCCTGCTCTGGCCCTTGCTATGATGATGACGAGCGAAGGAGCAATGTC	424
QY	61	CysLeuGlnCysLysThrLysThrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
DB	425	TGCCCTCAGTCAAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTTCATGGTGT	484
QY	81	GluGlyAspAspThrAspAlaasp---GlySerAspPheAsnTyrProAlaSerGly	99
DB	485	GAT---GAGGAGGAAGTGTGATGACCTGGCAATGAATTCACATAT	529
QY	100	ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly	119
DB	530	-----AAGCAA-----	535
QY	120	GlySerGlyAsnValGlyHisProLysTyr-----AspSerGlyGluIleGly	135
DB	536	-----GCCAATGGGAAGGCGCCAGAGTGGCAGCTTCAAGGAGATGACGCTCATCTGTCT	589
QY	136	LeuSer---LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsn	154
DB	590	TCACTGCTGCCATGAC-----CCACACCATCGGATTCACGCTTACAAAGT	637
QY	155	---SerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis---MetMetSer	172
DB	638	GGACAACAGATATCTGGAGAGATCCCTGATGCATCCCTGACCCGTCATCTATCCGCACT	697
QY	173	ProThrGlyAsn---IleSerArgArgAlaProPheProTyrValAsnHisSerProAsn	191
DB	698	CCAACTCGAGCTATGTTGATCCCAAGCGTTCAGTTCCTGTG-----AGGATTGTGGAC	751
QY	192	ProSerArgGluPhe---SerGlySerIleGlyAsnValAlaIleTrpLysGluArgValasp	210
DB	752	CCCTCGAGAGCTTGAATTCCTATGGGCTTAATAGTGTGATCGAAGGAAGAGTTGAG	811
QY	211	GlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAla	230
DB	812	AGCTGGAGGGTTAAACAGGACAAATAATGTTGCAAGTGAATAAT	856
QY	231	ProSerGluGlyArgAlaIleThrAspIleAspAlaSerThrGluTyrAsnMetGluasp	250
DB	857	---AAATATCCAGAGGCTAGAGGAGACATGGAG---GGGACTGGCTCAAATGGAGAAGAT	910

QY	251	AlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIleAlaSerSer	270
DB	911	ATGCAAAATGGTTGATGATGACGCGCTACCTTTTGGCGCATTTGTGCCAAATTCCTCAAAC	970
QY	271	LysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIlePheLeu	290
DB	971	CAGCTCAACCTTTACCGGATAGTAATCATCTCCGTCTTATCATCTCGTCTCTCTCTTC	1030
QY	291	HisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIle	310
DB	1031	CAATATCGTATCAGTCCATCCAGTGGCTAATGCTTATGGATTGGGTATCTGTTATC	1090
QY	311	CysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIle	330
DB	1091	TGTGAGTCTGTTGCTTGTCTTGTGTCAGTACAGTGGATCCATTGAAGGAACCTCCACTG	1150
QY	331	AsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSer	350
DB	1151	AACCGTGAGACATATCTCGACAGGCTTGCAITGAGGTATGATAGAGGAGGAGCCATCA	1210
QY	351	GlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProIle	370
DB	1211	CAGCTGGCTCCCATTTGATGCTTGTGTCAGTACAGTGGATCCATTGAAGGAACCTCCACTG	1270
QY	371	ValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSer	390
DB	1271	ATCAGAGCAACACACTGTTTGTCCATTTCTGCTGGATTACCTGTTGCACAAAGTGTC	1330
QY	391	CysTyrValSerAspGlyAlaSerMetMetLeuThrPheAspAlaLeuAlaGluThrSer	410
DB	1331	TGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGTAGTCTCTCTCTGAACTGC	1390
QY	411	GluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArgAlaPro	430
DB	1391	GAAATTTGCTAGAAAGTGGTTCCTTTTGAAGAAGCACAAATATTGAACCAAGAGCTCCA	1450
QY	431	GluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSerPheVal	450
DB	1451	GAAATTTACTTTTGCTCAAAAAATAGATTACCTGAAGGACAAAAATTCACCTTCATTTGT	1510
QY	451	LysAspArgArgAlaMetLysArgGluThrGluGluPheLysIleArgIleAsnAlaLeu	470
DB	1511	AAGGAAGACGAGCAATGAAGAGAGATGATGAAGAAATTCAAAAATAAGAAATCAATGCCTT	1570
QY	471	ValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrp	490
DB	1571	GTTGCCAAAGCACAAAGTGCCTGAGAGGGGTGGACCATGGCTGTTCTTGGGGCACAGTGT	1630
QY	491	ProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGly	510
DB	1631	CCTGGGAATAACCTTAGGACCATCTGGCATGATTCAAGTGTCTTCTTGGGGCACAGTGT	1690
QY	511	GlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArg	530
DB	1691	GGGCTTGACACTGATGAAATGAATACCACGCTTGTCTATGTCTCTCTCGTGAAGAGAGA	1750
QY	531	ProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAla	550
DB	1751	CCAGCTTTCAGCATCACAGAGGCTGTGTCATGATGATGATGATGATGATGATGATGATG	1810
QY	551	ValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAspSer	570
DB	1811	GTGCTGCAAAATGGTGCCTATCTCTCAATGTGATTTGTGACCATTTACTTCAATAGCAGC	1870
QY	571	LysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCys	590
DB	1871	AAAGCTTCTAGAGAAAGCAATGTGCTTCATGATGGATCCAGCTCTAGGAAGGAAACTGT	1930
QY	591	TyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArg	610
DB	1931	TATGTACAAATTTCCAAAGATTGTGTCATTTGACCTTGCCAGCATGATGCTAATAGG	1990

Qy	611	AsnThrValPhePheAspLeuArgGlyLeuAspGlyValLeuGlnGlyProValTyr	630	Db	3062	GAGCTATATGTTCAAGTGGACCAAGTTTGTCTCATCCCTCCGACCACTGTTCTTGTCAAT	3121
Db	1991	AACATAGTCTTCTTTGATATCAACATGAAGAGTCTAGATGGATTTCAGGTCATCTCTAT	2050	Qy	989	AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrp	1008
Qy	631	ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLys	650	Db	3122	AACTGGTGGGATGTTGGCAGGAATTCGTATGCATTAAACAGCGGTACCAATCTCTGG	3181
Db	2051	GTGGGAAACAGGATGCTGTTCATAGGCAGGCTTTGTATGGATATGATCTCTTTTGACT	2110	Qy	1009	GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe	1028
Qy	651	AlaLys-----LysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLysAlaSer	668	Db	3182	GGTCCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTGATCCTCATCTACCCCTTC	3241
Db	2111	GAAGCTGATCTGGAACCTAACATTGTTTAAAGAGCTGCTGGTAGAGGAAG---AGA	2167	Qy	1029	LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTrpAlaVal	1048
Qy	669	LysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSerVal	688	Db	3242	CTCAAGGCTCTCATGGGAGGAGAACCCGACCAACANTGCTATCTGTTGGTCCATC	3301
Db	2168	AGAACAAGAGGTTATATGATAGTCAAGCCGTTATATGAAGAACAGAACTTCAGCT	2227	Qy	1049	LeuLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProPhe-----Thr	1065
Qy	689	ProValPheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGlu	708	Db	3302	CTCTCTGCTCTATCTCTCTCTGCTGGGTGAGATCGATCTTCATCTCTCCCGACA	3361
Db	2228	CCCATCTTTAAATGAAGACATCGAGGAGGTTATTGAA-----GCTTATGAGGATGAA	2281	Qy	1066	ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyLysLeuAsnCys	1080
Qy	709	LysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPhe	728	Db	3362	CAGAAAGCTGCCGCC-----TTGGGGCAATGTGTGTAACACTGC	3400
Db	2282	AGGTCAAGTCTTATGTCCTCCAGAGAAATTTGGAGAAACGCTTTGTCAGTCTCCATCTTC	2341	RESULT 13			
Qy	729	ValAlaSerThrLeuMetGluTyrGlyValProGlnSerSerThrProGluSerLeu	748	AAZ99530			
Db	2342	ATTGCATCCACCTTTATGACTCAAGGTGGGATACACCTTCAACAAACCCAGCTTCTCTA	2401	AAZ99530 standard; DNA; 3753 BP.			
Qy	749	LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThr	768	AAZ99530;			
Db	2402	CTGAAGGAGCTATCATGTTATCAGCTGTGGGTACGAGGACAAATCTGAATGGGAA	2461	03-JUL-2000 (first entry)			
Qy	769	GluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHis	788	DNA encoding a maize cellulose synthase.			
Db	2462	GAGATTGGCTGGATCTATGTTTACAGAGGATATCTGACTGGGTITTAATAATGCAT	2521	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;			
Qy	789	AlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla	808	transgenic plant; plant breeding marker; ss.			
Db	2522	GCAAGAGCTGGCAATCACTACTGATGCCACCAGACCTGTTTCAAGGGTCTTGCA	2581	Zea mays.			
Qy	809	ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGlu	828	Key		Location/Qualifiers	
Db	2582	CCATCAATCTTCTGATGCTCTTAAATCAGGTGCTCGGTGGGCTCTTGGTCAGTGAA	2641	CDS	184..3405	/*tag= a	
Qy	829	IleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeu	848	FT	/product= "cellulose synthase"		
Db	2642	ATTCTGCTTAGCAGACATGCTCTATATGTTATGCTATGCTTACATGGCGGATGAACTTTG	2701	FT	/transl_except= (pos: 916..918, aa: Xaa)		
Qy	849	GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuVal	868	FT	/note= "no termination codon given; Xaa is an unspecified amino acid"		
Db	2702	GAGAGGCTGGCTTACATTACACCATGTTTATCCATCATCTGTTCCGCTTATCGCC	2761	PN			
Qy	869	TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSer	888	XX	W0200009706-A2.		
Db	2762	TATTTGTGCTCTGCTATCTGCTCTTACCAATAAAATTTATCATCTCTGAGATTAGT	2821	XX	24-FEB-2000.		
Qy	889	AsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu	908	XX	16-AUG-1999; 99WO-US018760.		
Db	2822	AAATATGCTGAATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2881	XX	17-AUG-1998; 98US-0096822P.		
Qy	909	GluMetArgTrpSerGlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpVal	928	XX	(PTON-) PIONEER HI-BRED INT INC.		
Db	2882	GAGCTCAGATGAGTGTGTGGCATTCGAAGATTGTGGAGAAATGACAGATTTTGGGT	2941	PI	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;		
Qy	929	IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAla	948	DR	WPI; 2000-224343/19.		
Db	2942	ATTGTTGGCAGCTCTGCGCATCTCTTCGCGGTGTTCCAGGGTCTGCTGAAAGTGTGCT	3001	XX	P-FSDB; AAY84120.		
Qy	949	GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAla	968	XX	New genes which encode maize cellulose synthase polypeptides in plants		
Db	3002	GGGATTGATACCACTTACAGTTACCTCAAGGATCTGATGAGATGGGACATTTGCT	3061	CC	useful for modulating the expression of cellulose synthase in plants and		
Qy	969	GluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIle	988	CC	to produce transgenic plants expressing the novel protein.		
				PS	Claim 1; Page 184-188; 119pp; English.		
				CC	The present sequence encodes a maize cellulose synthase polypeptide. The		
				CC	cellulose synthase can be used for the improvement of stalk quality for		
				CC	improved stand or silage. It also provides an increased concentration of		
				CC	cellulose in the pericarp, hardening the kernel and improving its		
				CC	handling ability. The sequences are used to produce transgenic plants and		
				CC	seeds expressing the cellulose synthase. The polynucleotide is used for		

Db	832	CAGGCAAAATATGATGCAGTGACTAAT-----AAATATCCAGAGGTAGA	879
Qy	236	AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp	255
Db	880	GGAGGA--GACATGGAG--GGGACTGGCTCAAATGGAGAAATATGCAAAATGGTTGAT	933
Qy	256	GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr	275
Db	934	GATGCACGGCTACCTTTGAGCGGTATCGTGCCAATTCTCAACACGAGCTCAACCTTAC	993
Qy	276	ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr	295
Db	994	CGGTAGTAGTCATCATTCGCTCTATCATCTCTGTCTCTCTTCTCCAGTATCGTGCAGT	1053
Qy	296	AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe	315
Db	1054	CATCCAGTGGTGATGCTTATGGATATATGGCTAGTATCTGTATCTGGAGGTCTGGTTT	1113
Qy	316	AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnAtqGluThrTyr	335
Db	1114	GCCTTGCTTGGCTCTAGATCAGTCTCCAAANAATGGTATCAATCAACCGTAGACATAT	1173
Qy	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaVal	355
Db	1174	CTTGACAGGCTTGCAATGAGGTATGATAGAGGGAGAGGCCATCACAGCTGGCTCCCATT	1233
Qy	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
Db	1234	GATGTCCTTCGTGAGTACAGTGGATCCATTTGAAGGAACTCCACTGTATCAGACCACACT	1293
Qy	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
Db	1294	GTITGTTCATCTCTTCGTGGATTAACCTGTTGACAAAGTGTCATGCTATGTTCTGAT	1353
Qy	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaAtqLys	415
Db	1354	GATGGTTCCAGTATGCTGACTTTTGTAGTCTCTCTCAGAAACCGCAGNAITTTGCTAGAAAG	1413
Qy	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
Db	1414	TGGGTTCCTCTTTGTAAGAACACAATATTGCAACAGAGCTCCAGAAATTTTACTTTGCT	1473
Qy	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
Db	1474	CAAAAATAGATTACCTGAAGGACAAATAATCAACCTTCATTGTTAAGAAAGACGCGCA	1533
Qy	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
Db	1534	ATGAAGAGGCGATGAGAAATTCAAAGTPAAGAAATCAATGCCCTTGTGCCAAGCACAG	1593
Qy	476	LysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
Db	1594	AAAGTGCCTGAAGAGGGGTGCACCATGGCTGATGGAACATGCATGGCTCGGAAATAATCT	1653
Qy	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1654	AGGACCATCTCGCATGATTACAGGTTCCTTGGGGCACAGTGGTGGGCTCGACACTGAT	1713
Qy	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATGAGTTACACGCTTTGTCTATGTCTCTCGTGAAGAGACACAGGCTTCAGCAT	1773
Qy	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAAGAAGCGCTGGTGCAATGAATGGCTGATTCGTGTATCTGTCTGCTGACAAATGGT	1833
Qy	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTCTCAATGTGGATTGCGACATTACTTCAATAGCAGCAAGCTCTTAGAGAA	1893
Qy	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

QY	296	AsnProValArgAsnAlaTyrProLeuTrrPheLeuSerValIleCysGluIleTrrPhe	315
Db	1054	CAATCAGTCGGTATGATATGATATATGGCTAGTATCTTATCTGCGAGCTTGGTTT	1113
QY	316	AlaLeuSerTrrPheLeuAspGlnPheProLysTrrPheProIleAsnAtcGluThrTyr	335
Db	1114	GCCTTGTCTGGCTCTAGATCAGTTCCTCAAAATGGTATCCAATCAACCGTAGACATAT	1173
QY	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaVal	355
Db	1174	CTTGACAGGCTTGCAATTCAGGTATGATAGAGAGGAGAGCCATCACAGCTGGCTCCCAT	1233
QY	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
Db	1234	GATGCTTCCTGTCAGTACAGTCGATCGATCAATGAAGAGCACTCCACATGATCAGACCA	1293
QY	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
Db	1294	GTTTTGTCCATCTTCTGTGGATTACCTGTGTGACAAAGTGTCATGCTATGTTCTTGAT	1353
QY	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
Db	1354	GATGGTTCAGCTATGCTGACTTTTGAGTCTCTCTCAGAAACCCAGAAATTTGCTAGAAG	1413
QY	416	TrrValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
Db	1414	TGGGTTCCTTTGTAGAGACCAATATTGAACCAAGAGCTCCAGAATTTTACTTTGCT	1473
QY	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
Db	1474	CAAAAAATAGATTACCTGAAGGCAAAATTCACCTTCATTGTTAAGNAAGACCGCGCA	1533
QY	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
Db	1534	ATGAAGAGGCGAGTATGAAGAATTCAAAGTAAGAATCAATGCCCTTGTGCCAAGCACAG	1593
QY	476	LysValProGluGluGlyTrrPheMetGlnAspGlyThrProTrrProGlyAsnAsnThr	495
Db	1594	AAAGTGCTCGAAGGGGTGGACCATGGCTGATGGAACTGCGTGGCTCGGGAATTAATCT	1653
QY	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1654	AGGGACCATCCTGGCATGATTGAGTTCCTGGGSCACAGTGGTGGGCTCGACACTGAT	1713
QY	516	GlyAsnGluLeuProArgIleValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATAGATTACCGTCTTGTCTATGTCTCTCGTGAAGAGACACAGGCTTTCCAGCAT	1773
QY	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAAGAGGCTGGTGCAATGAATGCGCTGATTCCTGTGATCTGCTGCTGACAAATGGT	1833
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTrrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGTGGATTGGACCATTAATCAATAGCAGCAAGCTCTTAGAGAA	1893
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

1054	Db	CATCCAGTCGGTATGATCTTATGGATATATGGCTAGTATCTTATCTGCGAGCTTGGTGT	1113
316	QY	AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnAtgGluThrTyr	335
1114	Db	GCCTTGCTTGGCTTCTAGATCAGTTCCTCAAAATGGTATCCATCAACCGTAGACATAT	1173
336	QY	LeuAspArgLeuAlaLeuArgTyrAspArgGluGluProSerGlnLeuAlaAlaVal	355
1174	Db	CTTGACAGGCTTGCAATGAGGTATGATAGAGAGGAGAGCCATCACAGCTGGCTCCCAT	1233
356	QY	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
1234	Db	GATGTCTTCGTCAGTACAGTCGATCCATTAAGAGAGCTCCACATGATCAGACCAAC	1293
376	QY	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
1294	Db	GTATTGTCCATCTTCTCTGTGATTACCTGTGTGACAAAGTGTCATCTATGTTTCTGAT	1353
396	QY	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
1354	Db	GATGTTTCAAGTATGCTGACATTTTGAGTCTCTCTCAGAAACCCAGAAATTTGCTAGA	1413
416	QY	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
1414	Db	TGGGTTCCCTTTGTAGAGACACATATTGAACCAAGAGCTCCAGAAATTTTACTTTGCT	1473
436	QY	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
1474	Db	CAAAAAATAGATATACCTGAAGGCAAAATTCACCTTCATTGCTTAAGGAAAGCCGCG	1533
456	QY	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
1534	Db	ATGAAGAGGAGATATGAAGAAATTCAAAGTAAGAAATCAATGCCCCCTGTGTCGCAAG	1593
476	QY	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
1594	Db	AAATGGCTCAAGAGGGGTGGACCATGGCTGATGGAACCTGCATGGCCCTGGGAATTA	1653
496	QY	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
1654	Db	AGGGACCATCCTGGCATGATTACAGGTTTCTTGGGGCACAGTGGTGGGCTCGACACTG	1713
516	QY	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
1714	Db	GGAAATAGATTACCGCTCTTGCTATGTCTCTCGTGAAGAGACACAGGCTTTCAGCAT	1773
536	QY	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
1774	Db	CACAAGAGGCTGGTGCAATGAATCGGCTGATTCTGTGATCTGCTGCTGACAAATGGT	1833
556	QY	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
1834	Db	GCCTATCTTCTCAATGTGGATTGGACCATTAATCAATAGCAGCAAGCTCTTAGAGAA	1893
576	QY	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

QY	316	AlaLeuSerTrpIleIleuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr	335
Db	1114	GCCTTGCTTGCTGCTTCTAGATCAGTTCCTCCAAATGGTATCAATCAACCGTGAGACATAT	1173
QY	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal	355
Db	1174	CTTGACAGGCTTGCAATTCAGGTATGATAGAGGAGGAGCCATCACAGCTGGCTCCCAT	1233
QY	356	AspIlePheValSerThrValAspProIleuLysGluProIleValThrAlaAsnThr	375
Db	1234	GATGTCTTCGTGTCAGTACAGTCGATCCATTTGAAGCAACCTCCACATGATCAGACCAACT	1293
QY	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
Db	1294	CTTTTGTCATCTTCTCTGGATTACCTCTTGACAAAGTGTCATGCTATGTTCTGAT	1353
QY	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
Db	1354	GATGGTTCCAGTATGCTGACTTTTGAGTCTCTCTCAGAAACCGAGNAATTTGCTAGAAAG	1413
QY	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
Db	1414	TGGGTTCCTTTTGTAAGAAGCAATATTTGAACAGAGCTCCAGAAATTTTACTTTGCT	1473
QY	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgArgAla	455
Db	1474	CAAAAAATAGNTTACCTGAGGACAAATTCACCTTCATTTGTTAAGGAAAGACGCGCA	1533
QY	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
Db	1534	ATGAAGAGGAGGATGAAGAATTCAAAGTAAGAATCAATGCCCTTGTGCCAAGCACAG	1593
QY	476	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
Db	1594	AAAGTGCTGAGAGAGGGTGACCATGTGTGTAAGAACTGCAATGGCTGGGAATATCTCT	1653
QY	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1654	AGGACCATCTGGCATGATTTCAGGTTTCTTGGGGCACAGTGTGGGCTCGACACTGAT	1713
QY	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATAGATTACCAAGCTCTGTCTATGTCTCTCGTGAAGAGAGACCAGGCTTTCCAGCAT	1773
QY	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAGAGGCTGGTGCAATGAATGCGCTGATTTCTGTATCTGCTGCTGCTGCACAAATGCT	1833
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGTGGATTGGACCATTAATTCATAGCAGCAAGCTCTTAGAGAA	1893
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

1174	DB	CTTGACAGCCTTGCAATGAGGTATGATAGAGGAGAGCCATCACAGCTGGCTGCCCAT	1233
356	QY	AspIlePheValSerThrValAspProLeuLysGluProPheIleValThrAlaAsnThr	375
1234	DB	GTGCTTCCTCGTCAGTACAGTGGATCCATTGAAGGAACCTCCACTGATCAAGCCAACT	1293
376	QY	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
1294	DB	GTITTGTCCTCATCTTCTGTGGATTACCTGTTGACAAAGTGTCACTGCTATGTTCTGAT	1353
396	QY	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
1354	DB	GATGGTTCAGCTATGCTGACTTTTGAGTCTCTCAGAAACCGAGAAATTGCTAGAAAG	1413
416	QY	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
1414	DB	TGGGTTCCTCTTTGTAGAAAGCACAATATTGAACAAAGAGCTCCAGAAATTTTACTTGCT	1473
436	QY	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
1474	DB	CAAAAATAGATTACCTGAGGACAAAAATTCACCTTCATTGTTAAGGAAGACGCGCA	1533
456	QY	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
1534	DB	ATGAAGAGGAGGATGAAGAATTCAAAGTAAGAATCAATGCCCTTGTCGCAAGACACAG	1593
476	QY	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
1594	DB	AAAGTGCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATATCTCT	1653
496	QY	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
1654	DB	AGGGACCATCTGGCATGATTACAGGTTTTCTTGGGGCACAGTGTGGGCTCGACACTGAT	1713
516	QY	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
1714	DB	GGAAATGAGTTACCAAGCTCTGTGTATGTCTCTCGTGAAGAGAGACAGGCTTTTCAGCAT	1773
536	QY	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
1774	DB	CACAGAGCGCTGGTGCAATGAATGGCTGATTTCTGTATCTTCGTGCTGCTGCACAAATGGT	1833
556	QY	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
1834	DB	GCCTATCTTCTCAATGTGGATTGGACCATTAATCAATAGCAGCAAGCTCTTAGAGAA	1893
576	QY	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

QY	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
DB	1234	GATGCTTCGTCAGTACAGTGGATCCATTGAAGGAACTCCACTGATCAGACCAACACT	1293
QY	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
DB	1294	GTITGTGCCATCTCTCTGTGGATTAACCTGTTGACAAAGTGTCATGCTATGTTCTGAT	1353
QY	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
DB	1354	GATGGTTTCAGTCATGCTGACTTTTGAGTCTCTCTCAGAAACCCGAGAAATTTGCTAGAAAG	1413
QY	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
DB	1414	TGGGTTCCCTTTGTGAAGACCAATATTGAACCAAGAGCTCCAGAAATTTTACTTTGCT	1473
QY	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
DB	1474	CAAAAAATAGATTACCTGAAGGACMAATTCAACCTTCATTGTTTGAAGAAACGCGCA	1533
QY	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
DB	1534	ATGAAGAGGGAGTATGAAGAATTCAAAGTAAGAATCAATGCCCTGTGTGCCAAGCACAG	1593
QY	476	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnIleThr	495
DB	1594	AAAGTGCTGAAGAGGGGTGACCATCGCTGATGGAACCTGCAATGGCGCTGGGAATAATCT	1653
QY	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
DB	1654	AGGGACCATCCTGGCATGATTTCAGGTTTCTTGGGGCACAGTGTGGGCTGGACACTGAT	1713
QY	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
DB	1714	GGAAATAGATTACCGCTCTGTCTATGTCTCTCGTGAAGAGAGACCAGGCTTTCCAGCAT	1773
QY	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
DB	1774	CACAAGAAGCGTGGTGCAATGAATGGCGTGAATTCTGTATCTGTCTGCTGCACAAATGCT	1833
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
DB	1834	GCCTATCTTCTCAATGTGGATTGGCACTTACTTCAATAGCAGCAAGACTCTTAGAGAA	1893
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

1234	GATGCTTCGTCAGTACAGTGGATCCATTGAAGGAACCTCCACTGATCAGACCAACACT	1293
376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
1294	GTTTTGTCATCTCTTCTGCGATTACCCCTGTTGACAAAGTGCATGCTATGTTCTGAT	1353
396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
1354	GATGCTTCAGCTATGCTGATCTTTGAGTCTCTCTCAGAAACCCGACGAATTTGCTAGAAAG	1413
416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
1414	TGGGTTCCTTTTGAAGAGCACAATATTGAACCAAGAGCTCCAGAATTTTACTTTGCT	1473
436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
1474	CAAAAAATAGATTACCTGAAGGACAAATTCACCTTCATTCTGTAAGGAAAGACGGCA	1533
456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
1534	ATGAAGAGGCGAGTATGAAGAATTCAAAGTAGAATCAATGCCCTCTGTGCCAAGCACAG	1593
476	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyValSerAsnThr	495
1594	AAAGTGCTTCAAGAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATTAATCCT	1653
496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
1654	AGGGACCATCTGGCATGATTACAGGTTTCTTGGGACACAGTGGTGGCTCGACACTGAT	1713
516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
1714	GGAAATGAGTTACCAAGTCTGTCTATGTCCTCGTGAAGAGACACAGGCTTTCCAGCAT	1773
536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValIleuThrAsnGly	555
1774	CACAAGAAGCTGGTGCAATGAATGGCGTGAATCTGTGTCATCTGCTGCTGACAAATGGT	1833
556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
1834	GCCTATCTTCTCAATGTGGATTGGCACCATTACTTCAATAGCAGCAAAAGCTCTTAGAGAA	1893
576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
1294	GTTTTGTCATCTTTCTGTGGATTACCCCTGTGACAAAGTGTCATGCTATGTTCTGAT	1353
396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
1354	GATGGTTTCAGCTATGCTGACTTTTGAGTCTCTCTCAGAAACCCGAGAAATTTGCTAGAAAG	1413
416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
1414	TGGGTTCCTTTGTAAAGACCAATATTGAACCAAGAGCTCCAGAATTTTACTTTGCT	1473
436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
1474	CAAAAAATAGATTACCTGAAGGCAAAATCAACCTTCATTCTTAAAGNAAGACGGCA	1533
456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
1534	ATGAAGAGCGAGTATGAAGAATTCAAAGTAAGAATCAATGCCCTGTGTGCCAAGCACAG	1593
476	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr	495
1594	AAAGTGCTGTAAGAGGGGTGCACCATGGCTGATGGAATGCATGGCTCGGGAAATATCTCT	1653
496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
1654	AGGACCATCTCGCATGATTACAGGTTTCTTGGGGCACAGTGGTGGGCTGCACACTGAT	1713
516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
1714	GGAAATGAGTTACCATGCTTGTGTATGTCCTCGTGAAAGAGACCCAGGCTTTCAGCAT	1773
536	HisLysIleAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
1774	CACAGAAGCGCTGGTGCAATGAATCGCTGATTCGTGTATCTGTGCTGTGTCGACAAATCGT	1833
556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgLys	575
1834	GCCTATCTTCTCAATGTGGATTGGACCATTACTTCAATAGCAGCAAGCTCTTAGAGA	1893
576	AlaMetCysPheIleuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

D	b		1294	GTTTGTCCATCTCTTCGTGGATTACCTGTTGCAGAAAGTGTCATGCTAATGTTCTGAT	1353
Q	y		336	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaLeuGluThrSerGluPheAlaArgLys	415
D	b		1354	GATGGTTCAAGTATGCTGACTTTTGAGTCTCTCTCAGAAACCAGAAATTTGCTAGNAAG	1413
Q	y		416	TrpValProPheValLysLysTyArgPheValProArgAlaProGluPheTyPheCys	435
D	b		1414	TGGGTTCCTTTTGTAAGAAGCAACAATATTGAACAGAGCTCCAGAAATTTTACTTTGCT	1473
Q	y		436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
D	b		1474	CAAAAATAGATTACTTGAAGGACAAAATACTCAACCTTCATTGTTAAGGAAGACCGCGCA	1533
Q	y		456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
D	b		1534	ATGAAGAGGCGATGAAGAAATTCAAATGATGAATCAATGCCCTTGTTGCCAAGCACAG	1593
Q	y		476	LysValProGluGluGlyThrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
D	b		1594	AATGTGCTGAAGAGGGGTGCACCATGGCTGATGGAATCGATGCGCTGGGAATAATCTCT	1653
Q	y		496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
D	b		1654	AGGACCATCTCGCANGATTACAGTTTTCTTGGGCGCAGTGTGTGGCTCGACACTGAT	1713
Q	y		516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
D	b		1714	GGAAATGAGTTACACGCTTGTCTATGTCTCTCGTGAAAGAGACCAAGCGCTTCAGCAT	1773
Q	y		536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
D	b		1774	CACAAGAAGCGCTGGTGCAATGAATGGCTGATTCGTGTATCTGTCTGCTGCACAAATCGT	1833
Q	y		556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
D	b		1834	GCCTATCTTCTCAATGTGGATTGCGACATTACTTCAATAGCAGCAAAGCTCTTAGAGAA	1893
Q	y		576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

[illegible]

1354	DB	GATGGTTCAGCTATGCTGACTTTTGTAGTCTCTCTCAGAAACCGCAGAAATTGCTAGAAAG	1413
416	QY	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
1414	DB	TGGGTTCCTCTTTGTAAAGAACCAATATTCACCAAGAGCTCCAGAAATTTTACTTTGCT	1473
436	QY	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
1474	DB	CAAAAATAGATTACCTGAGGACAAATATCAACCTTCATTGTGTAAAGGAACGCGCA	1533
456	QY	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
1534	DB	ATGAAGAGGCGATGTAAGAATTCAAAGTAAGAATCAATGCCCTTGTGCCAAGACACAG	1593
476	QY	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
1594	DB	AAAGTGCTGAAAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATAATCCT	1653
496	QY	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
1654	DB	AGGGACCATCTGGCATGATTACAGGTTTTCTTGGGGCACAGTGTGGGCTCGACACTGAT	1713
516	QY	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
1714	DB	GGAAATAGATTACCAAGCTCTTGTCTATGTCTCTCGTGAAGACAGACAGGCTTTCCAGCAT	1773
536	QY	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
1774	DB	CACAGAAGCGTGGTGCAATGAATGGCTGATTCGTGTATCTTCGTCTGCTGCACAAATGGT	1833
556	QY	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
1834	DB	GCCTATCTTCTCAATGTGGATTGGACCATTAATTCATAGCAGCAAGACTCTTAGAGAA	1893
576	QY	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

Qy	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
Db	1414	TGGGTTCCCTTTGTAGAAGACCAATATTGAACCAAGAGCTCCAGAATTTTACTTTGCT	1473
Qy	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
Db	1474	CAAAAAATAGATTACCTGAAGGACAAATTCAACCTTCATTGTTAAGAAACGCCGA	1533
Qy	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
Db	1534	ATGAAGAGGGAGGATGAAGAATTCAAAGTAAGAATCAATGCCCTTGTGCCAAGCACAG	1593
Qy	476	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
Db	1594	AAAGTGCTGAAGAGGGTGGACCATGGCTGATGGAACTGCGATGGCGCTGGGAATAATCT	1653
Qy	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1654	AGGGACCATCTGGCATGATTTCAGGTTTCTTGGGGCACAGTGTGGGCTCGACACTGAT	1713
Qy	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATAGATTACCGCTCTTGCTATGTCTCTCGTGAAAGAGACCAGGCTTTCCAGCAT	1773
Qy	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAAGAGCGCTGGTGCAATGAATCGCGTGATTCTGTGTTATCTGCTGCTGCACAAATGGT	1833
Qy	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGTGGATTGGACCACTTCTCAATAGCAGCAAAAGCTCTTAGAGAA	1893
Qy	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

1414	TGGGTTCCTTTTGTGAAGAGCAATATTGAACCAAGAGCTCGAATTTTACTTTGCT	1473
QY	GlnIysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
1474	CAAAAAATAGATTACCTGAAAGGCAAAATTCACCTTCATTCTTAAGGAAAGACGGCA	1533
QY	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
1534	ATGAAGAGGAGATGAAGAATTCAAAGTAGAATCAATGCCCTCTGTGCCAAGACACAG	1593
QY	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyIleAsnAsnThr	495
1594	AAAGTGCTCAAGAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATTAATCCT	1653
QY	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
1654	AGGAGCATCTCTGGCATGATTACAGGTTTCTTGGGCGCACAGTGTGGGCTCGACACTGAT	1713
QY	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
1714	GGAAATGAGTTACCAAGTCTGTGTATGTCTCTCGTGAAAGAGACACAGGCTTCAGCAT	1773
QY	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValIleuThrAsnGly	555
1774	CACAAGAGGCTGGTGCAATGAATGGCTGATCTGTGTATCTGTCTGTCTGCTACAAATGGT	1833
QY	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
1834	GCCTATCTTCTCAATGTGGATTGGCACTTACTTCAATAGCAGCAAAAGCTCTTAGAGAA	1893
QY	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

436	Gln	Lys	Ile	Asp	Tyr	Leu	Leu	Asp	Lys	Val	Gln	Pro	Ser	Phe	Val	Lys	Asp	Arg	Ala	455								
1474	CA	AAA	AA	TAG	AT	TAC	CT	GAA	GGA	CA	AA	AT	CA	ACC	TT	CAT	TT	CT	TAA	GA	CA	CG	CA	1533				
456	Met	Lys	Arg	Glu	Tyr	Glu	Glu	Phe	Lys	Ile	Arg	Ile	Asn	Ala	Leu	Val	Ser	Lys	Ala	Leu	475							
1534	AT	GAA	GAG	GAG	GAT	GAG	AA	TTC	AA	GA	AT	TTC	CA	AT	GCC	CT	TGT	GCC	AAG	CAC	AG	1593						
476	Lys	Val	Pro	Glu	Glu	Gly	Thr	Phe	Met	Gln	Asp	Gly	Thr	Pro	Trp	Pro	Gly	Asn	Asn	Thr	495							
1594	AA	AG	TGC	CT	GA	A	G	G	G	G	G	TG	CA	CT	GA	TG	GA	AA	CT	GC	CT	GG	GA	TA	TC	CT	1653	
496	Arg	Asp	His	Pro	Gly	Met	Ile	Gln	Val	Phe	Leu	Gly	His	Ser	Gly	Leu	Asp	Thr	Glu	515								
1654	AG	G	A	C	A	C	A	T	C	T	G	C	A	T	G	A	T	T	C	A	G	1713						
516	Gly	Asn	Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro	Gly	Phe	Cln	His	535							
1714	GGA	AA	TAG	ATT	ACC	AG	CT	CT	TGT	AT	GT	CT	TC	GT	G	AA	A	G	A	G	1773							
536	His	Lys	Lys	Ala	Gly	Ala	Met	Asn	Ala	Leu	Val	Arg	Val	Ser	Ala	Val	Leu	Thr	Asn	Gly	555							
1774	CAC	AA	GA	AG	G	CG	CT	GA	TGA	AT	GG	CT	GA	TTC	GT	AT	CT	GT	CT	GT	CT	GT	CT	GT	CT	GT	CT	1833
556	Gln	Tyr	Met	Leu	Asn	Leu	Asp	Cys	Asp	His	Tyr	Ile	Asn	Asn	Ser	Lys	Ala	Val	Arg	Glu	575							

1474	CAAAAAATAGATTACCTGAAGGCMAAATTCAACCTTTCATTTGTTAGGAAAGACGGCA	1533
QY	MetIysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerIysAlaLeu	475
DB	ATGAAGAGGCGATGAAGAATTCAAAGTAAGAATCAATGCCCTTGTGCCAAGACACAG	1593
QY	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
DB	AAAGTGCTCTGAAGAGGGGTGCACCATGGCTGATGGAACTGCATGGCTGGGAATAATCTCT	1653
QY	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
DB	AGGCACCATCTGGCATGATTACAGGTTTCTTGGGGCACAGTGGTGGGCTGCACACTGAT	1713
QY	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLeuArgProGlyPheGlnHis	535
DB	GGAAATGAGTTACCAAGCTCTTGCTATGTCTCTCGTGAAAGAGACACAGGCTTTCAGCAT	1773
QY	HisIysIleAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
DB	CACAGAAGCGCTGGTGCATGAATGGCGTGTATCGTGATCTGCTGTGCTGACAAATGCT	1833
QY	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerIysAlaValArgGlu	575
DB	GCCTATCTTCTCAATGTGGATTGGCACCATTACTTCAATAGCAGCAAAAGCTCTTAGAGAA	1893
QY	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

QY	456	MetLysArgGluTyrGluGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
DB	1534	ATGAAGAGGCGAGTATGAAGAATTCAAAGTAAGAATCAATGCCCTGTGTGCCAAGCACAG	1593
QY	476	LysValProGluGluGluTyrIleMetGlnAspGlyThrProTrrProGlyAsnAsnThr	495
DB	1594	AAAGTGCTGAAGAGGGGTGGACCATGGCTGATGGAACCTGCATGGCGCTGGGAATAATCTCT	1653
QY	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
DB	1654	AGGACCATCTGGCATGATTAGGGTTTTCTTGGGCGCACAGTGGTGGGCTCGACACTAGT	1713
QY	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
DB	1714	GGAAATGAGTTACCACGCTCTGTCTATGTCCTCGTGAAAGAGACACAGGCTTTCCAGCAT	1773
QY	536	HisGlyGlyAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
DB	1774	CACAAGAAGCGCTGGTGCATGAATGGCTGATTCTGGTGATCTGGCTGTCTGCAGCAATCGT	1833
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
DB	1834	GCCTATCTTCTCAATGTGGATTGGCACCATTACTTCAATAGCAGCAAAAGCTCTTAGAGAA	1893
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

Db	1534	ATGAAGAGGCGATGATGAAGAATTCAAGTAAGAAATCAATGCCCTTGTTGGCAAGCACAG	1593
Qy	476	LysValProGluGluGlyThrPleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
Db	1594	AAAGTGCTCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATAATCTCT	1653
Qy	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1654	AGGGACCATCTGGCATGATTACAGGTTTTCTTGGGGCACAGTGGTGGGCTCGACACTGAT	1713
Qy	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATGAGTTACACAGCTCTTGCTATGTCTCTGCTGAAAAGAGACCAAGGCTTTCAGCAT	1773
Qy	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAGAGCGCTGGTGCATGAATGCGCTGATTCTGCTGTATCTGCTGTGCTGCACAAATGCT	1833
Qy	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGTGGATTGGACCATTACTTCAATAGCAGCAAGAGCTCTTAGAGAA	1893
Qy	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

Qy	476	LysValProGluGluGlyTyrPheMetGlnAspGlyThrProTyrProGlyAsnAsnThr	495
Db	1594	AAAGTGCTGAAGAGGGGTGACCATGGCTGATGGAACTGCATGGCTCGGGAAATAATCTCT	1653
Qy	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1654	AGGGACCATCTGGCATGATTTCAGGTTTCTGGGGCAGAGTGTGGGCTGGACACTGAT	1713
Qy	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATGAGTTACCAAGCTCTGTGTATGTCTCTGCTGAAAGAGACCAGGCTTTTCAGCAT	1773
Qy	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAAGAGGCTGGTGCAATGAATGGCTGATTTCTGTCTATCTGTCTGCACAAATGGT	1833
Qy	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGTGGATTGGACCATTAATTCATAGCAGCAAGAGCTCTTAGAGAA	1893
Qy	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

1594	AAAGTGCTGAAGAGGGGTGGACCATGGCTGATGGAGCACTGCATGGCTGGGAAATATCT	1653
Db		
496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Qy		
1654	AGGACCATCTGGCATGATTGAGGTTTCTTGGGGCACAGTGGTGGCTCGACACTGAT	1713
Db		
516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Qy		
1714	GGAATAGATTACCATCGTCTGTGTCATGTCTCTCGTGAAAGAGACCAAGGCTTTCAGCAT	1773
Db		
536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Qy		
1774	CACAAGAAGGCTGGTGCAATGAATGGCTGATTCGTGTATCTGCTGCTGACAAATCGT	1833
Db		
556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Qy		
1834	GCCTATCTTCTCAATGTGGATTGGACCATTACTTCAATAGACCAAGCTCTTAGAGAA	1893
Db		
576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595
Qy		

[illegible]

Db	1654	AGGGACCATCTGGCATGATT	CAGGTTTCTTGGGGCAGATGGTGGGCTGACACTAGT	1713
QY	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis		535
Db	1714	GGAAATGAGTTACCATCGTCTTGCTATGTCTCTCGTGAAAGAGACCCAGGCTTTTCAGCAT		1773
QY	536	HisLysGlyAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly		555
Db	1774	CACAAGAAGCGTGGTGCAATGAATGCGCTGATTCTCGTGATCTGCTGCTGCTGACAAATCGT		1833
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu		575
Db	1834	GCCTATCTTCTCAATGTGGATTGCGACCACTACTTCAATAGCAGCAAAAGCTCTTAGAGAA		1893
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro		595

Qy	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1714	GGAAATAGATTACACGCTCTTGTCATGTCTCTCGTGAAGAGACCAAGGCTTTCCAGCAT	1773
Qy	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAAAGAGCGCTGGTGCAATGAATCGGTGATTCTGCTGTCTGTCGACAAATCGT	1833
Qy	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGGATTGGCACTTACTTCAATAGCAGCAAGCTCTTAGAGAA	1893
Qy	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

1714	Db	GGAAATGAGTTACCAAGTCTTGTCATGTCCTCGTGAAGAGACACGAGCTTTCCGCAAT	1773
536	QY	HisIvsiVsalacIyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
1774	Db	CACAAAGCGCTGGTGCAATGAATCGCTGATTCGTGTATCTGCTGTGTCACAAATCGT	1833
556	QY	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
1834	Db	GCCTATCTTCTCAATGTGGATTGGACCATTACTTCAATAGCAGCAAGCTCTTAGAGAA	1893
576	QY	AlaMetCysPheIeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

Qy	536	HisLysLysAlaGluValMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1774	CACAGAGCGTGGTGCATGAATGCGCTGATTCGTGTATCTGCTGCTGCACAAATGCT	1833
Qy	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1834	GCCTATCTTCTCAATGTGGATTGCGACCATCTTCAATAGCAGCAAGCTCTTAGAGAA	1893
Qy	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

D	b	1774	CACAGAAAGCGTGGTGCAATGAATGCCTGATTCTGTATCTGCTGTCTGACAATGGT	1833
Q	y	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
D	b	1834	GCCTACTTTCTCAATGTGGATTGCGACATTACTTCATAGCAGCAAAGCTCTTAGAAA	1893
Q	y	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595

[illegible]

Db 1834 GCCTATCTTCTCAATGTGATTCGACCATCTACTTCAATAGCGCAAGCTCTTAGAGAA 1893

Qy 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595

QY 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595

Db	1894	GCAATGCTCTCATGATGGATCCGGCTCTAGGAAGAAAACCTTGTTATGTACAAATTTCCA	1953	QY	954	PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe	973
QY	596	GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe	615	Db	3025	TTTCAGATTACCTCAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTC	3084
Db	1954	CAGAGATTGATGGCATTGACCTTGACAGATCGATATGCTTAATCGGAACATAGCTTTCTTT	2013	QY	974	LysTrpThrThrLeuLeuLeuProProThrThrThrThrLeuLeuLeuLeuValGlyVal	993
QY	616	AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys	635	Db	3085	AAAGTGAGCCAGTTGCTCATCTCCGACCATGTTCTTGTTCATTAACCTGGTCCGAATG	3144
Db	2014	GATATCAACATGAAAGGCTGGATGGCAATTCAGGGTCCAGTTTACGTGGGAACAGATGC	2073	QY	994	ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly	1013
QY	636	ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----Lys	653	Db	3145	GTGGCAGGAAATTTCTTATGCCAATTAACAGTGGCTACCAATCTCTGGGGTCCGCTCTTTGA	3204
Db	2074	TGTTTCAATAGACAGCTTTGATGATAGATCGATCTCTTTGACTGAAGCTGATCTGGAG	2133	QY	1014	LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet	1033
QY	654	ProGlyPheLeuAlaSerLysCysGlyLysLysLysLysLysLysLysLysLysLys	673	Db	3205	AAAGTCTTCTTCGATCTGGGTGATCTCTCCATCTCTACCCCTTCTCAAGGGTCTCATG	3264
Db	2134	CMAACATGTATTAAAGAGCTGCTGTGTAGAAAGAG-----AAAAAGAACAAAGATTAT	2190	QY	1034	GlyArgGlnAsnArgThrProThrThrThrThrThrThrThrThrThrThrThrThr	1053
QY	674	SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu	693	Db	3265	GGAAAGCAGAACCGCACACCAACATCGTCATTTGTCTGTCCATCTCTTGCATCTATC	3324
Db	2191	ATGGATAGTCAAAAGCCGTTATTATGAAGAGAACAGAAATCTTCAGTCCCATCTTCAATATG	2250	QY	1054	PheSerLeuLeuTrpValArgValAspProPhe-----ThrThrArgLeuAlaGly	1070
QY	694	GluAspIleGluGluValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet	713	Db	3325	TTCTCTTCTGCTGGGTGAAGATCGATCTTTTCATCTCCCGCACACAGAAAGCTGCTGCC	3384
Db	2251	GAAAGACATCAAGAGGGTATTGAA-----GGTTACGAGGATGAAGGTCAGTGTCTATG	2304	QY	1071	ProAsnIleGlnThrCysGlyIleAsnCys	1080
QY	714	SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu	733	Db	3385	-----TTGGGCAATGTGGCGTCAACTGC	3408
Db	2305	TCCAGAGGAAATTTGGAGAAACGCTTGGTCAGTCTCTTATTTTCAFTGATCCACCTTT	2364	RESULT 14			
QY	734	MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle	753	ID	AAZ99515	standard; DNA; 3753 BP.	
Db	2365	ATGACACAAGGTGGCATACCCACTTCAACAAACCCAGCTTCTTCTACATAAGGAGCTATC	2424	XX	AAZ99515;		
QY	754	HisValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIle	773	XX	03-JUL-2000	(first entry)	
Db	2425	CATGTCATCAGTTGTGATATCAGGACAAACTGAATGGGAAAGAGATTGGCTGATC	2484	XX	DNA encoding a maize cellulose synthase.		
QY	774	TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArg	793	XX	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;		
Db	2485	TATGTTTCAGTAACGGAGGATATTCTGACTGGGTTTAAATGTCATGCAAGGGCTGGCAA	2544	XX	transgenic plant; plant breeding marker; ss.		
QY	794	SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer	813	OS	Zea mays.		
Db	2545	TCAATCTACTGATGCCACCGACCTTGTTCAGGGTTTCAAGGGTTCTGCACCAATCACTTCC	2604	XX	Key	Location/Qualifiers	
QY	814	AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg	833	XX	CDS	184..3405	
Db	2605	GATGCTCTTAATCAGTGTCTCGTTGGCTCTTGGGTGAGTGAATTCGTTAGTAGA	2664	FT		/*tag= a	
QY	834	HisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr	853	FT		/product= "cellulose synthase"	
Db	2665	CATTGCTCTATCTGTTGTTTACAATGACGATTTGAAGCTTTTGGAGGGCTGGCTTAC	2724	FT		/trans except= (pos: 916..918, aa: Xaa)	
QY	854	IleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuPro	873	FT		/note= "no termination codon given; Xaa is an unspecified	
Db	2725	ATCAACACTAATGTATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	2784	XX		amino acid"	
QY	874	AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle	893	XX			
Db	2785	GCTATCTGCTCTTCTTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2844	XX			
QY	894	TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer	913	XX			
Db	2845	TTCTTCTCTCTTCTTCTGCTCTTCTGCTCTTCTGCTCTTCTGCTCTTCTGCTCTTCT	2904	XX			
QY	914	GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSer	933	XX			
Db	2905	GGTGTGGGATTAAGAGGATGGTGGGAAATGACAGATTTGGGTATTGGTGGCCTCT	2964	XX			
QY	934	AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn	953	XX			
Db	2965	GCCCATCTCTTCCAGTGTTTCCAGGGTCTGCTGAAAGTGTGGCTGGGATTGATCAAC	3024	XX			

New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein.

Claim 1; Page 144-149; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monoots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Sequence 3753 BP; 944 A; 805 C; 965 G; 1038 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	0	Length:	3753
Score:	3981.00	Matches:	747
Percent Similarity:	80.64%	Conservative:	132
Best Local Similarity:	68.53%	Mismatches:	163
Query Match:	68.90%	Indels:	48
DB:	3	Gaps:	19

US-09-900-237A-30 (1-1080) x AAZ99515 (1-3753)

QY	2	AspGlyAspAla---AspAlaLeuIysSerGlyArgHisgIyAlaGlyAspValCysGln	20
Db	250	GACGGCGATGTCGGGGCTCGCTAAGCCCAAGAGTGGAAATGGACAGGTCTGCCAG	309
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
Db	310	ATTTCGGTGATCTGTGGGTGTTTCAGCCACTGGTGATGCTTTGTGTGCTGCAATGAG	369
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla	60
Db	370	TGTGCTCTCCCTGTCTCGCGCCATGCTATCAGTATGAGCGCAGAGGGGGAACCAATGC	429
QY	61	CysLeuGlnCysLysThrLysTyrIlyArgHisArgGlySerProAlaIleArgGlyGlu	80
Db	430	TGCCCCCACTGCAAGACTAGATACAAGAGACAGAAAGGTAGCCCTCGAGTTCATGCTGAT	489
QY	81	GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly	99
Db	490	GAG---GATGAGGAGAGTGTGATGACCTAGACAATGAATTCACATAC	534
QY	100	ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly	119
Db	535	-----NAGCAA-----	540
QY	120	GlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyr	139
Db	541	GGCAGTGGGAAGGCCCAAGCTGGCAGCTACCTCAAGGAGATGATCTGATCTGCTTCATCT	600
QY	140	AspSerGlyGluIleProArgGlyTyrValProSerValThrAsn---SerGlnMetSer	158
Db	601	GCTGCCCATGAG---CCACATCATCGGATTCACGGCTTGACAGCGCGTCAACAGATATCT	657
QY	159	GlyGluIleProGlyAlaSerProAspHisHis---MetMetSerProThrGlyAsn---	176
Db	658	GGAGAGATTCCTGATGCTCTCCCTGACCGTCACTTATCCGCAGTCCACATCGAGCTAT	717
QY	177	IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe	196
Db	718	GTTGATCCACGGCTCCAGTTCCTGTG-----AGGATTTGTCGCCCTCGAAGACATTG	771

QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575	Db	2905	GGTGTGGCATTGAAGATTGGTGGAGAAATGACAGTTTGGGTATTGGTGGACCTCT	2964
Db	1834	GCCTATCTCTCAATGTGATTCGCGACCAATCTCAATAGCAGCAAGCTCTTAGAGAA	1893	QY	934	AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValleAlaGlyIleAspThrAsn	953
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595	Db	2965	GCCCATCTCTTCGACGTTCGAGGCTCTGCTGAAAGTGTGGCTGGGATTGATCAAC	3024
Db	1894	GCAATGTGCTTCATGATGATCCGCTCTAGGAGGAAACCTGTGTATGACAAATTC	1953	QY	954	PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe	973
QY	596	GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe	615	Db	3025	TTTCAAGATTACCTCAAGGCACTGATGAGGATGGCGACTTGTGCTGAGCTATATGTGTT	3084
Db	1954	CAGAGATTGATGGCATTGACCTTGCAGATCGATATGCTAATCGGAACATAGTTTCTTT	2013	QY	974	LysTyrThrThrLeuLeuLeuProProThrThrIleLeuIleIleAsnMetValGlyVal	993
QY	616	AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys	635	Db	3085	AAGTGCACCAAGTTTGTCTTCCTCCGACCACTGTTCTTGTCAATTAACCTGGTGGAAAT	3144
Db	2014	GATATCAACATGAAGGCTCTGGATGTCATTCAGGCTCCAGTTTACGTGGGAACAGATGC	2073	QY	994	ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly	1013
QY	636	ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----Lys	653	Db	3145	GTGCGAGGAATTTCTTATGCCATTAACAGTGGCTACCAATCTTGGGTCGCTCTTTGGA	3204
Db	2074	TGTTTCAATAGACAGGCTTGTATGATACGATCTCTTGTGACTGAAGCTGATCTGGAG	2133	QY	1014	LysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMet	1033
QY	654	ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysLysArg	673	Db	3205	AAGCTGTCTCTCGATCTGGGTGATCCTCCTTCTACCCCTTCTCAAGGCTCTCATG	3264
Db	2134	CCAAACATTGTTTAAAGAGCTGCTGTGTAGAGGAAG---AAAAGAAACAAGATTAT	2190	QY	1034	GlyArgGlnAsnArgThrProThrIleValIleValIleValTyrAlaValleLeuAlaSerIle	1053
QY	674	SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu	693	Db	3265	GGAAGGCAGAACCGCACCAACAATCGTCAATGTCTGTGGTCCATCTCTTGTGATCTATC	3324
Db	2191	ATGGATAGTCAAAAGCCGCTATTATGAAGAGAACAGAACTCTCAGCTCCCTCAATATG	2250	QY	1054	PheSerLeuLeuTyrValArgValAspProPhe-----ThrThrArgLeuAlaGly	1070
QY	694	GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet	713	Db	3325	TTTCTCTTGTCTGGGTGAAGTGCATCTTTTCTATCTCCCGACACAGAAAGCTGCTGCC	3384
Db	2251	GAAGACATCGAAGGGTATTGAA-----GGTTACGAGGATGAAGCTCAGTCTTATG	2304	QY	1071	ProAsnIleGlnThrCysGlyIleAsnCys	1080
QY	714	SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaPheValAlaSerThrLeu	733	Db	3385	-----TTGGGCAATGTGGCTCAACTGC	3408
Db	2305	TCCAGAGGAATTTGGAGAACGCTTGTGTCAGTCTCTATTTTCATTGCTCCACCTTT	2364	RESULT 15			
QY	734	MetGluTyrGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle	753	AAZ99497			
Db	2365	ATGACACAGGTGGCATACCACTTCAACAAACCAGCTTCTCTACTAAGGAAGCTATC	2424	ID	AAZ99497	standard; DNA; 3780 BP.	
QY	754	HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle	773	XX	AAZ99497;		
Db	2425	CATGTCATCAGTTGTGATATGAGGACAAACTGAATGGGGAAGAGATTGGCTGATC	2484	XX	03-JUL-2000	(first entry)	
QY	774	TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg	793	DT		DNA encoding a maize cellulose synthase.	
Db	2485	TATGGTTCAAGTACGAGGATATCTGACTGGGTTTAAATGTCATGCAAGCGGTGGCAA	2544	DE		Maize; cellulose synthase; stalk quality; stand; silage; cellulose;	
QY	794	SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer	813	KW		transgenic plant; plant breeding marker; ss.	
Db	2545	TCAATCTACTGATGCCACCGACCTTGTTCAGGGTTTCTGACCAATCAATCTTTC	2604	XX			
QY	814	AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg	833	OS		Zea mays.	
Db	2605	GATCGCTTAATCAGTCTCGTGGGCTCTTGGTCACTGGAAATCTGCTTAGTAGA	2664	XX		Location/Qualifiers	
QY	834	HisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr	853	XX		201..3422	
Db	2665	CATTGTCTCTATCTGTATGGTTTACAAATGACAGCTTGTGGAGGGCTGCTTAC	2724	FT		/*tag= a	
QY	854	IleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuPro	873	FT		/product= "cellulose synthase"	
Db	2725	ATCAACACTATTGTATATCCAAATCAATCCATCCGCTTATGCTTATGTGCTTCC	2784	FT		/trans_except= (pos: 933..935, aa: Xaa)	
QY	874	AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle	893	FT		/note= "no termination codon given; Xaa is an unspecified	
Db	2785	GCTATCTGCTCTTACCAATAAATTAATCATCTCTGAGATTAGCAATATGCTGGATG	2844	XX		amino acid"	
QY	894	TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer	913	PN		WO200009706-A2.	
Db	2845	TTCTTCAATCTCTTTTTCGCTCCATTTTTCGCTGATATGAGCTTAGATGGAGT	2904	XX		24-FEB-2000.	
QY	914	GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTyrValIleGlyIleSer	933	XX		16-AUG-1999; 99WO-US018760.	
				XX		17-AUG-1998; 98US-0096822P.	
				XX		(PION-) PIONEER HI-BRED INT INC.	
				XX		Dhugga KS, Helentjaris TG, Bowen BA, Wang X;	
				XX		WPI: 2000-224343/19.	
				XX		P-PSDB; AAY84109.	

PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.

PS Claim 1; Page 97-102; 119pp; English.

XX
 CC The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein

SQ Sequence 3780 BP; 946 A; 817 C; 976 G; 1040 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 0 Length: 3780
 Percent: 3981.00 Matches: 747
 Best Similarity: 80.64% Conservative: 132
 Best Local Similarity: 68.53% Mismatches: 163
 Query Match: 68.90% Indels: 48
 DB: 3 Gaps: 19

US-09-900-237A-30 (1-1080) x AA299497 (1-3780)

QY 2 AspGlyAspAla---AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 DB 267 GACGGGATGTCGGGGTCGGCTAAGCCCAAGAGTGCGAATGGACAGTCTGCCAG 326
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 327 ATTTCGGGTGACTCTGTGGGTGTTTCAGCCACTGGTGATGCTTTGTTCCTGCAATGAG 386
 QY 41 CysArgPheProValCysArgProCysTyroGluHisGluArgLysGluGlyThrGlnAla 60
 DB 387 TGTGCTTCCTGCTGTCGGCCCATGCTATGATGATGAGCAAGGAGGGAACCAATGC 446
 QY 61 CysLeuGlnCysLysThrLysTyroLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 447 TGCCCCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTCATGGTAT 506
 QY 81 GluGlyAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly 99
 DB 507 GAG---GATGAGGAAGATGTTGATGACTAGACAAATGAATCAACTAC----- 551
 QY 100 ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly 119
 DB 552 -----AAGCAA----- 557
 QY 120 GlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyr 139
 DB 558 GGCAGTGGGAAAGGCCAGAGTGCACAACTGCAAGGAGATGATGCTGCTCTTCTATCT 617
 QY 140 AspSerGlyGluIleProArgGlyTyrValProSerValThrAsn---SerGlnMetSer 158
 DB 618 GCTCGCCATGAG---CCACATCATCGGATTCACGGCTGACAGCGGTCAACAGATATCT 674
 QY 159 GlyGluIleProGlyAlaSerProAspHisHis---MetMetSerProThrGlyAsn--- 176

DB 675 GGAGAGATTCTCTGATGCTTCCCTGACCGTCATTTCTATCGCAGTCCCAACATCGAGTAT 734
 QY 177 IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe 196
 DB 735 GTTGATCCCAAGCGCTCCAGCTTCCTGTG-----AGGATTGTGACCCCTCGAAGGACTTG 788
 QY 197 ---SerGlySerIleGlyAsnValAlaTlpLysGluArgValAspGlyTlpLysMetLys 215
 DB 789 AATTCTCTATGGGCTTAATAGTGTGACTGGAAGAAAGAGTTGAGAGCTGAGAGGTTAAA 848
 QY 216 GlnAspGlyGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
 DB 849 CAGGACAAATAATGATGCAAGTCACTAAT-----AAATATCCAGAGGCTAGA 896
 QY 236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
 DB 897 GGAGGA---GACATGGAG---GGGACTGGCTCAATGGAGAAATATGCCAATGGTTGAT 950
 QY 256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
 DB 951 GATGCACGGCTACCTTTGAGCGGTATCGTGCCTAATTTCTCAACAGCTCAACCTTTAC 1010
 QY 276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
 DB 1011 CCGGTAGTATCATTTCTCGTCTTATCATCTGTCTTCTTCCAGTATCGTGTCACT 1070
 QY 296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe 315
 DB 1071 CATCCAGTGGGTGATGCTTATGCAATATGGCTAGTATCTGTTATCTGCGAGTCTGGTTT 1130
 QY 316 AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr 335
 DB 1131 GCCTTGTCTGGCTTCTAGATCAGTTCCCAAAATGGTATCCCAATCAACGGTGAGACATAT 1190
 QY 336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal 355
 DB 1191 CTTGACGGCTTGCAATTGAGGTATGATAGAGGGAGAGCCATCACAGCTGGCTCCCAT 1250
 QY 356 AspilePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375
 DB 1251 GATGCTTCGTGTCAGTACAGTGGATCCATTGAAGGAACCTCCACTGATCACAGCCAACT 1310
 QY 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395
 DB 1311 GTTTTGTCCATTCTTCTGTGATTACCTCTGTGACAAAGTGTCTGCTATGTTCTGAT 1370
 QY 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
 DB 1371 GATGGTTTCAGCTATGCTGACTTTTGTAGTCTCTCTCAGAAACCGCAGAAATTTGCTAGA 1430
 QY 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435
 DB 1431 TGGGTTCCCTTTGTAGAGAGCACATATTGAACCAAGAGCTCCAGAAATTTTACTTTGCT 1490
 QY 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
 DB 1491 CAAAAAATAGATTACCTGGAAGGACAAATTCACCTTCATTGTTAGGAAAGACGCGCA 1550
 QY 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
 DB 1551 ATGAAGGGGAGTATGAAGAATCAAGTAAGAAATCAATGCCCTTGTGTCGAAGCACAG 1610
 QY 476 LysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr 495
 DB 1611 AAGTGCCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATAATCT 1670
 QY 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
 DB 1671 AGGACCATCTGGCATGATTGAGGTTTCTTGGGCGCACAGTGGTGGCTCGACACTGAT 1730
 QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgLysArgProGlyPheGlnHis 535

Db	1731	GGAAATGAGTTACACGCTTGTCTATGCTCTCTCGTGAAGAGACACAGGCTTTTCAGCAT	1790
QY	536	HisLysIysAlaGlyValMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1791	CACAGAGGCTGGTGGCAATGAATCGCTGATTCGTGTATCTGCTGCTGACAAATGGT	1850
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1851	GCCTATCTTCAATGTGGATTGGACCATTAATCAATAGCAGAAAGCTCTTACAGAA	1910
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595
Db	1911	GCAATGTGCTTCAATGATGATCCGGCTCTAGAGAGAAACTTGTTATGTACAAATTTCCA	1970
QY	596	GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe	615
Db	1971	CAGAGATTGATGGCATTCACITTCGACGATCATATGCTTAATCGGAACATAGTTTCTTT	2030
QY	616	AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys	635
Db	2031	GATATCAACATGAAGGCTGGATGTCATTCAGGGTCCAGTTTACGTGGGAACAGATGC	2090
QY	636	ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----Lys	653
Db	2091	TGTTTCAATAGACAGGCTTTGATGATACGATCTCTTTTGACTGAAGTGATCTGGAG	2150
QY	654	ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLysSerLysLysArg	673
Db	2151	CCAAACATTGTTATTAGAGCTGCTGGTGGTAGAAGGAAG---AAAAAGAACAGAGATTAT	2207
QY	674	SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu	693
Db	2208	ATGGATAGTCAAAAGCCGTATTATGAAGAGAACAGAAATCTTCAGTCCCATCTTCAATATG	2267
QY	694	GluAspIleGluGluGluValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet	713
Db	2268	GAGACATCGAAGAGGGTATTGAA-----GGTTACGAGGATGAAGGTTCAGTGTATG	2321
QY	714	SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu	733
Db	2322	TCCAGAGGAAATGGAGAAACGGTTTGGTCAGTCTCTATTTTTCATTGATCCACCTTT	2381
QY	734	MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLysGluAlaIle	753
Db	2382	ATGACACAAGGTGGCATACCACTTCAACAAACCCAGCTTCTTACTTAAGAGAGCTATC	2441
QY	754	HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle	773
Db	2442	CATGTCATCAGTTGTGATATGAGACAAACAACTGAATGGGGAAAGAGATTGGCTGGATC	2501
QY	774	TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg	793
Db	2502	TATGGTTCACTAACGGAGGATATCTGACTGGGTTTAAATGCAATGCAAGGGGCTGGCAA	2561
QY	794	SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer	813
Db	2562	TCAATCTACTCATGCCACCAACGACCTTGTTCAGGGTTCTGACCAATCAAICTTTTCC	2621
QY	814	AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg	833
Db	2622	GATGCTTATATCAGGTGCTCCGTTGGGCTCTTGGTTCAGTGGAAATTTCTGCTTAGTAGA	2681
QY	834	HisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr	853
Db	2682	CATTGCTCATCTGATGTTGATACCAATGACGATTTGAAGCTTTTGGAGGGCTGGCTTAC	2741
QY	854	IleAsnThrThrIleTyrProLeuThrSerLeuLeuValTyrCysIleLeuPro	873
Db	2742	ATCAACACTATTGTATATCCCAATCAATCCATTCGCTTATTCGCTTATTTGTTGCTTCCC	2801
QY	874	AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle	893
Db	2802	GCTAICTGCTCTTACCAATAAATTTATCAATTCCTGAGATTAGCAATTTATGCTGGGATG	2861
QY	894	TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer	913
Db	2862	TTCITTCATTTCTTTTCGCCTCCATTTTTCGCCACTGGTATATTGGAGCTTAGATGGAGT	2921
QY	914	GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyIleSer	933
Db	2922	GGTGTGGCATGTGAAGATTGGTGGAGAAATGAGCAGATTTTGGGTATTGGTGGCACCTCT	2981
QY	934	AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn	953
Db	2982	GCCCATCTCTTCGCAAGTGTTCAGAGGTCTGCTGAAAGTGTTCGCTGGGATTGATACCAAC	3041
QY	954	PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe	973
Db	3042	TTTCAGATTACCTCAAAGGCATCTGATGAGGATGGCGACTTTTGCTGAGCTATATGTTTC	3101
QY	974	LysTyrThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal	993
Db	3102	AAAGTGGACCAAGTTTGTCTATTCTCCGACCATCTTAAACAGTGGCTACCAATCCTGGGTCCGCTCTTTGGA	3161
QY	994	ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly	1013
Db	3162	GTGGCAGGAAATTTCTATGCCATTAAACAGTGGCTACCAATCCTGGGTCCGCTCTTTGGA	3221
QY	1014	LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet	1033
Db	3222	AAAGTGTCTTCTCGATCTGGGTGATCTCCCATCTCTACCCCTTCTCAAGGGTCTCATG	3281
QY	1034	GlyArgGlnAsnArgThrProThrIleValIleValIleValIleValLeuAlaSerIle	1053
Db	3282	GGAAAGCAGAACCGCACCAACATCGTCAITGTCTGGTCCATCTCTTTCATCTATC	3341
QY	1054	PheSerLeuLeuTyrValArgValAspProPhe-----ThrThrArgLeuAlaGly	1070
Db	3342	TTCTCTTGTCTGGGTGAGATCGATCTTTCATCTCCCGCACACAGAAAGCTGCTGCC	3401
QY	1071	ProAsnIleGlnThrCysGlyIleAsnCys	1080
Db	3402	-----TTGGGCAATGTGGCGTCAACTGC	3425

Search completed: August 22, 2004, 17:53:11
Job time : 1096 secs